Social Psychology
The Second Edition

Roger Brown

FP

To Albert Gilman

6

Group Polarization

THE WELL-KNOWN TELEVISION PLAY and movie *Twelve Angry Men* (Rose, 1955) dramatizes the situation in which a single juror favors acquittal in a trial for first-degree murder and so faces a majority of eleven favoring conviction. The play opens with the judge's instructions to the jury in which he tells them that they are required to return a unanimous verdict. If, he stresses, there is "any reasonable doubt," the verdict must be "not guilty," but if there is no doubt, it is their duty to bring back a verdict of "guilty." The jury then moves into the jury room, and the whole action of the play is the deliberation that takes place there. An initial ballot is taken almost at once: Eleven vote "guilty"; one votes "not guilty." When the play ends, the majority has been converted to the minority position, and the unanimous verdict is "not guilty." Audiences at the movie version know from the very start that "not guilty" is the correct verdict, because the lone holdout is played by Henry Fonda. At the end, they also know it because what transpires in the jury room has convinced them.

Twelve Angry Men is treated as something of a joke in books on jury process (Saks and Hastie, 1978) because it is known from records of real trials (Kalven and Zeisel, 1966) and also from laboratory and computer simulations (Tanford and Penrod, 1983) that a switch from 11-1 in one direction to 12-0 in the other is an extreme statistical rarity, in fact, assigned zero probability in most prediction models. *Twelve Angry Men* is therefore, some think, an absurdly unrealistic play. The trouble with that conclusion is that audiences always find the play to-

tally convincing and so most presumably judge it to be psychologically realistic. On what basis can they do so?

Psychological realism in fiction, or truth in fiction, cannot mean typicality in life. It would be hard to make a drama of a jury process that moved from an initial ballot of 11-1 in one direction to a final verdict of 12-0 in the same direction. It would be true in the sense of typical, but would it make a play, especially a three-act play? Some surprise somewhere, something not typical, is essential to the arousal of interest. Perhaps the movie of *Twelve Angry Men* is atypical yet convincing because the lone holdout is Henry Fonda, from whom prodigies of heroic accomplishment are credible. In fact, however, that is not what makes the play "work." And it does work; it works like a perfectly designed mechanical toy. It works because the jury deliberations conform to known principles of group dynamics, which would in fact produce the decision reversal represented.

In general, I think realism in fiction must be judged against the implicit or tacit psychology that we all know. We must know a lot of psychology on the implicit level, because humans have from the beginning worked steadily, if informally, at figuring one another out, guessing at thoughts and feelings so as to anticipate actions. The routine of life could not proceed as smoothly as it does without a large amount of shared psychological knowledge, and I think we probably judge the truth of fiction against this criterion rather than the simple criterion of event probability. Sometimes when the events are very improbable, writers assure us, "This is a true story," but that assurance has become the very hallmark of a story writers have failed to make credible, as witness *The Amityville Horror*—a true story.

The unusual thing about *Twelve Angry Men*, and it may even be unique in this respect, is that we can judge it to be psychologically realistic against explicitly formulated psychological principles. It is written as if its author, Reginald Rose, had studied the topic of group polarization and tailored his play to the principles that explain it—which he could not have done since the play antedates the discovery of the principles. It will be my goal in this chapter to prove the realism of *Twelve Angry Men* by telling the story of group polarization.

The Risky Shift

We must begin at a point remote from the play and from jury deliberations or anything to do with the law, so it may be well first to expose the essential structure of *Twelve Angry Men*; on that level of abstraction we shall always be discussing the same thing. What happens in the play is, very abstractly, this. At time 1, designated individuals (in the play it is the jurors) arrive individually, in advance of any discussion among themselves, at decisions on a problem that has been presented to all (the trial proceedings in the play). At time 2, the indi-

viduals convene as a group and discuss the problem with the intention of reaching a collective decision. At time 3, the group decision is reached (the verdict in the play). In *Twelve Angry Men* interest focuses on the relation between the individual decisions at time 1 (the initial ballot), which were 11–1 in favor of guilty, and the group decision at time 3, the 12–0 verdict "not guilty." The theory of group polarization describes a regular relationship between the distribution of initial decisions and the eventual group decision and identifies two processes in the intervening discussion that are responsible for the relationship. Therefore, the remarkable reversal in the play is an instance of the kind of thing the theory must explain.

Prediscussion individual decisions on a problem and postdiscussion group decisions characterize the operations of every sort of board and committee. The decision may involve admission of an applicant to a medical school, approval of an application for a research grant, or, for that matter, the decision of President John Kennedy's advisers to make the Bay of Pigs invasion of Cuba or the decision of President Reagan's advisers to go ahead with the 1985 income tax revisions. If there is a law in this domain, it is likely to be consequential.

There are many dimensions on which decisions can vary and, therefore, many dimensions on which group decisions might consistently differ from individual decisions. One such dimension, one that was important in the decision about the Bay of Pigs, is *riskiness*. Committees and boards in business and government have been generally thought to favor the pussyfooting compromise rather than the bold action. Whyte (1956), for example, has argued that the team approach in business enterprises leads inexorably to an inhibition of daring and risk-taking, a preference for the conservative course when a choice must be made between more and less risky courses of action. The academic committees on which one has served, committees to decide on appointments, course offerings, and degree recommendations, do not leave one breathless at their daring. Consequently, the discovery by Stoner (1961) that group decisions following discussion were, on certain problems, consistently riskier than individual decisions immediately awakened great interest.

Stoner first reported the effect in a dissertation submitted to the School of Industrial Management at MIT. Stoner used problems that were devised by Wallach and Kogan (1959, 1961) and first used by them for the investigation of individual differences in risk-taking. Each problem describes a situation in which the central person is faced with a choice between two actions whose outcomes differ in their attractiveness and their probability. The respondent is to act as an adviser to the central person in the story and is asked to indicate the minimum probability of success that he would demand before recommending the alternative that might have the more desirable outcome. Here is an example of the problems used by Stoner:

1. Mr A, an electrical engineer who is married and has one child, has been working for a large electronics corporation since graduating from college five years ago. He is

assured of a lifetime job with a modest, though adequate, salary, and liberal pension benefits upon retirement. On the other hand, it is very unlikely that his salary will increase much before he retires. While attending a convention, Mr. A is offered a job with a small, newly founded company with a highly uncertain future. The new job would pay more to start and would offer the possibility of a share in the ownership if the company survived the competition of the larger firms.

Imagine that you are advising Mr. A. Listed below are several probabilities or odds of the new company's proving financially sound. Please check the *lowest* probability that you would consider acceptable to make it worthwhile for Mr. A to take the new job.

[   ] The chances are 1 in 10 that the company will prove financially sound.

[   ] The chances are 3 in 10 that the company will prove financially sound.

[   ] The chances are 5 in 10 that the company will prove financially sound.

[   ] The chances are 7 in 10 that the company will prove financially sound.

[   ] The chances are 9 in 10 that the company will prove financially sound.

[   ] Place a check here if you think Mr. A should *not* take the new job, no matter what the probabilities.

Brief résumés of several of the other problems, numbered as they were by Stoner, will suggest their range:

3. A man of moderate means may invest some money he inherited in secure "blue-chip" low-return securities or in more risky securities that offer the possibility of large gains.

4. A captain of a college football team, in the final seconds of a game with the college's traditional rival, may choose a play that is almost certain to produce a tie score, or a more risky play that would lead to sure victory if successful, sure defeat if not.

6. A college senior planning graduate work in chemistry may enter University X where, because of rigorous standards, only a fraction of the graduate students manage to receive the Ph.D., or he may enter University Y which has a poorer reputation but where almost every graduate student receives the Ph.D.

In this chapter I shall call the problems Stoner used by the name they came to be known by in social psychology: "Choice Dilemmas."

What is a "risky" decision, and how is the amount of risk determined? A person in a state of risk must have something to lose, a stake. For Mr. A in the first story above it is an assured lifetime job with an adequate salary. To take a risk is voluntarily to endanger that stake. Presumably neither Mr. A nor anyone else would do such a thing unless there were also a prize; the prize in the present case is a superior job in a new firm. If the individual is to have a problem of decision-making, the prize must exceed the value of the stake but be less certainly available than the stake. If the prize and stake were equally valuable and yet the probability of attaining the prize varied, as it does in problem 1 above, then Mr. A would not consider risking his stake. Would he give up his

present job in order to take a new job exactly like the one he has in a company that might fail? Probably not.

Stoner's subjects were male graduate students of industrial management. They first studied the problems, twelve problems in all, and made individual decisions on each problem. Subsequently, they were assembled in groups of six and instructed to discuss each problem and to arrive at a unanimous group decision. Twenty-three other subjects did not meet as groups but did study the problems a second time, after a lapse of a few weeks. Stoner put together thirteen groups, and for twelve of them the predominant direction of shift on the problems between the means of the initial individual decisions and the later group decisions was toward greater risk. The twenty-three control subjects showed no systematic shift in either direction.

Stoner also asked his subjects to record their private judgments after the group decision had been made; it was understood that a man's private opinion might or might not agree with the group consensus. Those private opinions, subsequent to discussion, were compared with the private opinions expressed in advance of discussion. About 45 percent of the subjects did not change their private views; of the remainder, however, 39 percent changed toward greater risk and only 16 percent toward greater caution. Something in the group discussion appears to have influenced private opinions, as well as the group decision, in the direction of greater riskiness. The change came to be called (grammar be damned) the "risky shift."

Stoner's finding proved not to be a one-time outcome; it was promptly replicated with several kinds of subject. Wallach, Kogan, and Bem (1962) repeated the procedure with more than two hundred undergraduate students in the liberal arts curriculum at the University of Colorado. Groups were either all male or all female. Fourteen out of fourteen male groups shifted in a risky direction, and twelve out of fourteen female groups did the same. In this experiment, as in Stoner's, subjects were asked to record their individual decisions following the discussion, and while their decisions often did not agree with those reached by the group, they ran to greater risk than the decisions made prior to the group discussion.

In 1971 a special issue of *The Journal of Personality and Social Psychology*, under the editorship of Dean Pruitt (1971a, 1971b), was devoted to papers on the risky shift. By that date, one decade after Stoner wrote his thesis, the effects he obtained with the Choice Dilemmas had been replicated so many times that people had stopped counting. One decade further along (Myers, 1982) it was estimated that more than a hundred studies in a dozen countries had been done with the Choice Dilemmas, and the original effects had always been obtained.

It is important to define precisely the risky shift that is so robust a result. It is not the case that every group that discusses a Choice Dilemma shifts to greater risk, and it is still less the case that every individual does. It is not the case that groups must be of the same size, six members, as those Stoner used; the risky shift has been reliably obtained with groups of two to seven members,

and probably larger groups would also work. It is not the case that the group convened to hold a discussion must be required to reach a consensus decision, because the risky shift occurs between initial individual opinions and final post-discussion opinions as well as between initial opinions and the group's consensual decision. The exact definition of the risky shift, the one that can be replicated at will, is: The mean risk scores on the Choice Dilemmas obtained from individuals prior to a group discussion are less risky than the mean of group scores (across groups) obtained after discussion.

A Sample Discussion

In our introductory social psychology course, we have for many years used the Choice Dilemmas and the full risky shift design (initial individual decisions—group discussion—group decision—final individual decisions) as a laboratory exercise. The exercise works beautifully, but one must be careful to forewarn a class that the risky shift does not occur with every group, but only as the mean across several groups, and that the effect is not large, but usually on the order of one unit.

From many discussions of problem 1 (Mr. A, the electronic engineer) I have created a composite discussion and imagined a characteristic risky shift outcome.

The participants are six men whom we shall identify by the letters A to F.

A: Let's see where we stand. I favor taking the job if the odds are one in ten.

B: Only if they are five in ten.

C: Right, five in ten.

D: I say seven in ten.

E: My judgment is three in ten.

F: Also three in ten.

A: (to D) Where do you get that seven in ten?

D: This guy is not on his own; he's not free to do whatever he—. He has a wife and child to support.

C: He will have the expense and trouble of moving, and he will lose his retirement benefits, which means something after all.

B: Why should he leave a sure thing, a perfectly secure job? Only greed could lead him to—.

A: What's the matter with you guys? Security! Retirement! What're you, half dead? Be dynamic! Go forward, have a spirit of adventure. Nothing ventured, nothing gained!

E: My reasoning is this: The man is out of school almost five years; he is a relatively young man, school-wise or experience-wise. I feel he has nothing to lose by taking the chance of going with a new, young company and

possibly going much farther than he will as an organization man in some huge outfit.

C: You're all talking as if this is the only chance he's ever going to get to move. Just because he turns down this job doesn't mean he has to stay put the rest of his life. Within two years or so he will get another offer, so why should he go with a company that only has one chance in ten of making it? Statistics on the failure and success of new businesses show that the large majority actually succeed.

F: Yes, but the most he can lose on this change is moving expenses. I read in *The New York Times* or somewhere that there is a big demand for electrical engineers. Even if this company should fall through, he can always get a job, probably at the same salary. I figure even if the company goes broke, he can still claim valuable experience in looking for another job later on. That's why I'd say three in ten.

C: That's true; if the company fails, it isn't a personal failure. I propose a compromise on three in ten.

D: I can go along with that.

B: O.K.

E: O.K.

A: Why not?

The initial individual decisions of the persons constituting this group were: 1 in 10, 3 in 10, 3 in 10, 5 in 10, 5 in 10, and 7 in 10. The mean of these values is 4 in 10, and the unanimous group decision is represented to be 3 in 10. Let us imagine that the subjects were also asked to record their personal decisions following the discussion and that it was pointed out to them that these might not agree in every case with the group consensus. In Figure 6-1, we have an imaginary but typical set of final positions compared with the initial positions. The two effects common to all the experiments we have cited can be observed in these data: The unanimous group decision is riskier than the mean of initial individual decisions; the mean of the final individual positions (3 in 10) is also riskier than the mean of initial positions. Notice, however, that it is not the case (see Figure 6-1) that every individual changes to a riskier decision; what happens is that those initially near the middle of the scale become more extreme in the risky direction.

The phenomenon called the "risky shift" was thought of as a surprising fact about the difference between individual decisions and group decisions. The effect has held up for twenty years, but the nature of the effect has turned out to be incorrectly conceptualized as a shift to greater risk; it must be more abstractly defined. However, the idea of the risky shift lasted long enough for people to propose two explanations for it, conformity or leadership, and for those explanations to fail (Brown, 1965). It is worth reviewing the usefulness of conformity and leadership, because they fail also to explain the reconceptualized phenomenon that replaced the risky shift.

Explanations That Fail

Because in the risky shift (see Figure 6-1) individual initial positions converge after discussion, the social force of conformity (Chapter 1) seems a likely explanatory principle. Certainly the individual judgments do converge after discussion, as Figure 6-1 shows. The convergence on a unanimous group decision is fully accounted for by the experimenter's instructions to reach such a decision, and one need not invoke conformity to explain that outcome; it looks more like simple obedience to authority. However, there is also substantial convergence in the individual final positions (the mean is, like the group decision, 3 in 10), and that individual convergence is explicitly *not* required by the experimenter. It probably represents genuine changes of opinion consequent to discussion. However, there is more to the risky shift than convergence or conformity.

If the individual decisions in Figure 6-1 converged to the exact mean of their initial positions (4 in 10), that would be a simple conformity effect. Or if very many groups were to discuss the problem of Mr. A and if some converged to the mean of their original position, some to positions somewhat riskier than
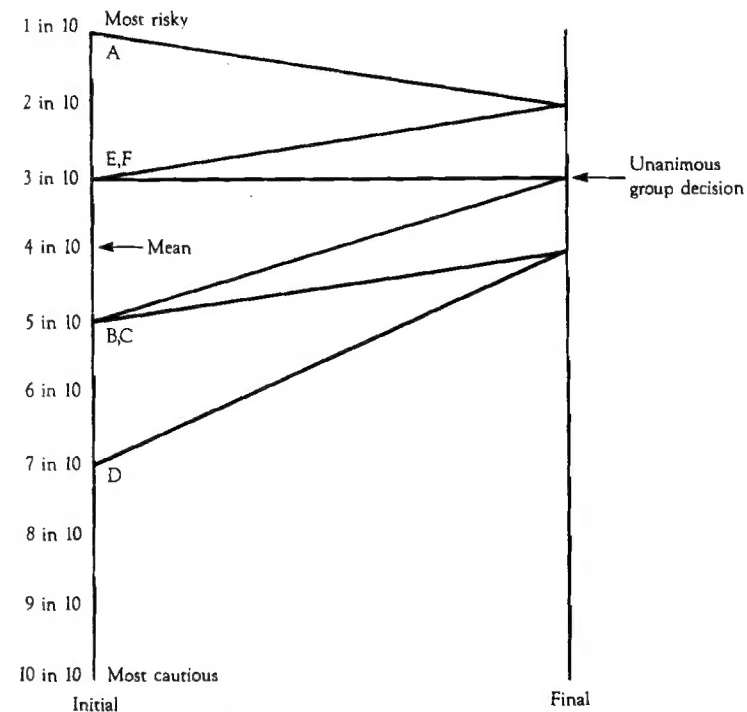

Figure 6-1. Initial positions, final positions, and the unanimous group decision on a problem involving risk

the mean, and some to positions somewhat more cautious, such that the overall mean of the final positions was the same as the mean of the initial positions, that too would be a simple conformity effect. However, the risky shift is not convergence to the mean of initial positions, but, rather, to points on one side of the mean, the riskier side. Therefore, conformity cannot account for the effect.

The decision of a group will settle on a position that is not the mean of the group if some members are disproportionately influential. When one member is notably more influential than the others, he is a leader. Like the leader of a group of nomadic anthropoids, he is the one "out front," the one who seems to select the way that all will go. Perhaps the shift to risk in group decisions on the Stoner problems is a phenomenon of leadership.

There is clear evidence that qualities of leadership do not account for the risky shift. In one study (Burnstein and Vinokur, 1973), after individuals had recorded their initial decisions, they were required to go into the group discussion and defend a position that was the mirror image of their real preference; e.g. if someone chose a risky 2 in 10, he had to argue for a cautious 8 in 10. If it were the case that general qualities of leadership (charisma?) characterized high risk takers, then those qualities should enable them to cause any position they defended to prevail. Since persons really disposed to high risk were defending cautious positions, one looked for a shift to caution, but it did not happen. On normally risky Choice Dilemmas, no shifts at all occurred. The influentiality of high risk takers does not export.

Several other kinds of result rule out leadership as the explanatory principle. It was shown many times in the 1970s that choice shifts could be produced without group discussion by simply having group members announce aloud their individual decisions, and there would seem to be no way to exercise leadership in such circumstances. In addition, it was discovered that shifts on the Choice Dilemmas were not always in the direction of greater risk but were sometimes in the direction of greater caution. That rules out an explanation of high risk-takers as leaders. However, it also rules out the risky shift as the phenomenon to be explained and is, in fact, the critical discovery that moved the entire discussion up one notch in generality.

Risky Shift or Cautious Shift

The fact is that the Choice Dilemmas had produced some cautious shifts from the very first (Stoner, 1961). A cautious shift is defined in exact parallel with a risky shift as a change in the mean of groups or of final individual postdiscussion decisions by contrast with the mean of initial individual decisions. For a short time it was possible to overlook the cautious shifts (they occurred on Stoner's problems No. 5 and No. 12), because results were summed across the full set of twelve Choice Dilemmas, and in the full set risky shifts swamped cautious shifts. When attention was first paid to the fact that some shifts were to

greater caution on the Choice Dilemmas, there was a moment of discomfort in which we thought we had lost our phenomenon. If collective decisions are sometimes riskier than individual decisions and sometimes more cautious and (it eventually was noticed) sometimes not significantly different, then there is no generalization at all, because all logical possibilities occur.

The concern was short-lived. What dispelled it was the realization that it was always the same problems (No. 5 and No. 12) on which cautious shifts occurred. With many studies of the Choice Dilemmas on record, it was possible to see that individual dilemmas tended to breed true; that is, always to produce the same sort of shift, either to risk or to caution. That meant there was a general phenomenon, there was a regularity, though it was more complex than had originally been thought. It took some years and a lot of work to find the best way to conceptualize the new regularity.

What kind of problem produces a shift after discussion to increased caution? Here is the first one that consistently did so, Stoner's No. 12.

Mr. M is contemplating marriage to Miss T, a girl whom he has known for a little more than a year. Recently, however, a number of arguments have occurred between them, suggesting some sharp differences of opinion in the way each views certain matters. Indeed, they decide to seek professional advice from a marriage counselor as to whether it would be wise for them to marry. On the basis of these meetings with a marriage counselor, they realize that a happy marriage, while possible, could not be assured. Listed below are several probabilities or odds that their marriage could prove to be a happy and successful one.

Imagine that you are advising Mr. M and Miss T. Please check the *lowest* possibility that you would consider acceptable for advising Mr. M and Miss T to get married.

Nordhøy (1962), who was the first to take a serious interest in cautious shifts, succeeded in writing some new choice dilemmas that consistently produced a shift to increased caution, as did David Myers, Colin Fraser, and others, and soon there was a good-sized stock of problems known to have as a reliable property the power to produce either the risky shift or the cautious shift. Here is a problem that was not included in Stoner's original set that produces a strong cautious shift:

A good friend of yours, call him Sam, is about to board a plane at the airport to begin his overseas vacation. He has been looking forward to this trip very much, but he is troubled because he awoke in the morning with a quite severe abdominal pain. Because Sam has never flown before, he thinks that the pain may simply be an upset stomach brought on by anticipation of the flight. Although he is not far from a hospital where he knows he could obtain quick attention, he realizes that a visit to the hospital would cause him to miss his flight which, in turn, would seriously disrupt his vacation plans. The pain seems to have grown more severe in the last few minutes. Listed below are several probabilities or odds that Sam's stomach trouble will go away.

Imagine that you are advising Sam. Please check the *lowest* probability or odds

that you would consider acceptable for Sam to go ahead with his trip. [Stoner, 1968, p. 446]

While no one has quite worked out a formula for writing the two kinds of problem to order, it is possible to identify one property of a story problem that makes caution likely: a very large stake. If someone's life or marriage is at stake, the cautious shift is likely. Another usual property of cautious shift problems is the involvement in the decision of others besides the protagonist. Caution is likely when a risky decision threatens a fiancée, a family, or parents, and not only the protagonist. As for what makes people disdain to be cautious, a small stake and very large prize help. Here, for instance, is a dilemma (Myers, 1982) with those qualities:

> Henry is a writer who is said to have considerable creative talent but who so far has been earning a comfortable living by writing cheap Westerns. Recently he has come up with an idea for a potentially significant novel. If it could be written and accepted, it might have considerable literary impact and be a big boost to his career. On the other hand, if he was not able to work out his idea or if the novel was a flop, he would have expended considerable time and energy without remuneration.
>
> Imagine that you are advising Henry. Please check the *lowest* probability that you would consider acceptable for Henry to attempt to write the novel.

The fact that some choice dilemmas reliably produce risky shifts and that other choice dilemmas reliably produce cautious shifts is a regularity, to be sure, but not a very useful one. What is wanted is a general characterization of the kinds of problem that will produce each effect so that one can predict beyond those that have been tried out and found to work. One sort of characterization is suggested by the "hints" listed above for writing items that make people cautious, but those are just hints and have not led to any interesting results. A shift predictor of a completely general kind has been found, and though it may at first seem less interesting than a content formula would be, its discovery has, in fact, helped to unlock the deeper reasons why shifts occur.

A Predictor of the Shifts

Teger and Pruitt (1967) were, I think, the first to identify the shift predictor. They used the twelve original Choice Dilemmas and the full standard design, defining the shift as the difference between the mean of initial decisions and the mean of final postdiscussion decisions. The new question they asked of their data was whether or not any general relationship existed between the mean of the initial decision and the size of the shifts that occurred. Looking first at just the ten problems that usually produced risky shifts (exclusive, that is, of Stoner's No. 5 and No. 12), it was apparent that there was considerable variation from item to item in the size of risky shift produced, from a tiny shift of .16 to a sizable shift of 1.48. The shift of .16 is not actually large enough to be signif-

icant and so counts as no shift at all, and that was true of a couple of other items, though most of the shifts were large enough to be significant. We know today, when so much work has been done with the Choice Dilemmas, that the tendency of items to "breed true" is not limited to the gross division between risky shift types and cautious shift types, but holds true on the level of the individual item. That is to say that among risky shift items and among cautious shift items, each one tends to hold a characteristic place in terms of both the direction and the size of the shift produced.

Returning to the item that produced no significant risky shift (.16) and the item that produced a healthy shift of 1.48, it turned out that there was a surprising difference in the mean of the initial decisions. The item that shifted a lot had a quite extreme risky initial position whereas the item that shifted a little or not at all started out near the middle of the scale. Across the ten risky shift items there was a significant correlation of −.64 between mean initial preference and mean shift. The correlation is negative because high risk means low odds, and the relationship was such that the riskier the initial disposition an item evoked (across all subjects), the larger the shift to increased risk. The two items that produced cautious shifts (No. 5 and No. 12) had initial positions that were less risky and so more cautious (odds of about 8 in 10 were the minimum acceptable), and the cautious shifts they produced may be thought of as negative risky shifts, and so the correlation can be computed across all twelve items, and it rises to −.78. The meaning of that correlation is the same for both risky and cautious items: a group discussion moves decisions to more extreme points *in the direction of the original inclination*. Group discussion produces polarization, which means shift to either risk or caution in the direction of the original disposition, and the size of shift increases with the degree of initial polarization.

There are many studies demonstrating that initial decisions on choice dilemmas predict sizes of shift on the same items, but the sharpest and clearest is that of Myers and Arenson (1972). They worked with six of the original Stoner problems, all of which reliably produced risky shifts, and with six new items written (successfully) to produce cautious shifts. Subjects were all females, and groups ranged in size from two to seven. The correlation between mean initial positions on items and shifts on the same items was −.89, which is very high indeed. The plot of initial means against mean shifts (Figure 6-2) is especially instructive. Notice first that shifts to greater caution are represented on the ordinate as negative risky shifts. There are just four cautious shifts, and they occur where the initial minimal odds acceptable are 7 in 10 or greater. If you follow the 0.0 shift line across, you will see that there is a region of neutrality on the odds scale which seems to extend (roughly) from about 5.5 to about 7. We can think of these data as indicating that items that initially dispose individuals to risk (defined as odds of 5.5 in 10 or lower) elicit more risky decisions after discussion whereas items that initially dispose to caution (defined as odds of 7 in 10 or greater) elicit more cautious decisions after discussion, and items occupying a psychologically neutral region (5.5 in 10 to 7 in 10) show no signifi-
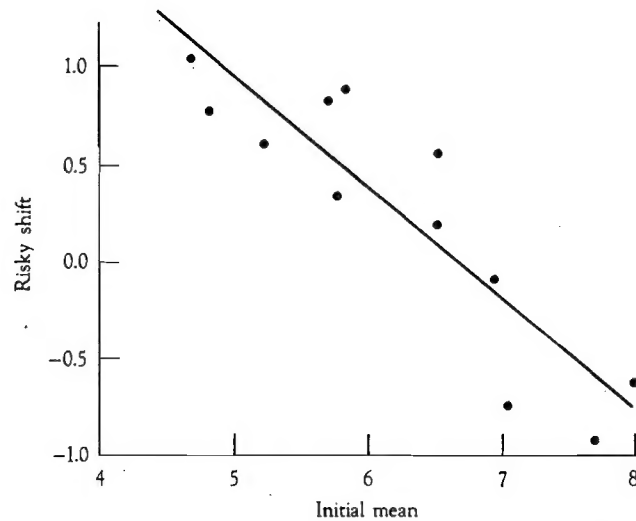
Figure 6-2. Initial means and mean shifts on twelve choice dilemmas

(Reprinted with permission of authors and publisher from D. G. Myers and S. J. Arenson, Enhancement of dominant risk tendencies in group discussion, *Psychological Reports* [1972], *30*:615–23, Fig. 1)

cant shift. The psychologically neutral region does not, as it happens, exactly coincide with the midpoint of the scale. These findings are very representative of the whole set of similar studies.

One aspect of the relationship pictured in Figure 6-2 can be puzzling. The relationship is such that the more extreme the original preference, the greater the shift in the same direction. One might think that extremity to begin with would limit or put a ceiling on further extremity. Notice, however, that the riskiest initial position in Figure 6-2 is not even as risky as odds of 4 in 10, whereas odds of 1 in 10 are possible; so for those data there is plenty of room for further polarization to be produced by discussion. There is also quite a bit of room at the cautious extreme for further polarization. The other relevant consideration, and it is very important to keep this in mind, is that the initial mean for an item is a *property of that item* across multiple groups. This means that a problem that averages odds of 4 in 10 across groups may, for any particular group, yield a less risky mean, such as 5 in 10 or 6 in 10. That is a fact that has a future in this discussion, because it might be taken to mean that the overall initial mean of an item is its real position on a risk–caution scale and a group scoring nearer the midpoint on that item may have a lot of room for change.

Explanation of the Shifts

There are two processes—"social comparison" and "persuasive arguments"— that account for the shifts in risk. In the original group discussions, the two act

together and reinforce each other. I think it is now clear that either process alone is sufficient to produce a significant shift (whether to increased risk or to increased caution). I also think it likely that the two processes are about equally powerful. In this evaluation of two decades of research, I believe I am in agreement with Fraser (1971), Myers and Lamm (1976), Lamm and Myers (1978), Pruitt (1971a, 1971b), Sanders and Baron (1977), and others. However, Eugene Burnstein (1982) and others believe that one of the two processes, "persuasive arguments," is alone necessary and sufficient to explain all findings.

The evidence supporting each process is by now quite massive and must be accorded proper respect. I myself, however, respond to something in addition to the evidence, which is a little click of intuitive rightness. Just what produces the click is hard to say; it has something to do with simplicity, something to do with introspective recognition, and something to do with tightness of fit among the several component parts of each process.

The phenomenon we have come to know as shifts *in* risk, that is shifts to either risk or caution, was for some years thought of as the risky shift, or shifts to increased risk only, and both the social comparison explanation and the persuasive arguments explanation date, in nascent form, from the risky shift period. That is to say that the two processes, especially social comparison, were originally devised to explain the risky shift. Later on it proved possible to generalize both processes so as to explain all shifts in risk, whether risky or cautious. However, a certain amount of early experimental evidence was collected to test explanations of the risky shift only, and so some individual studies are asymmetrical from our present vantage point. Eventually, however, completely parallel sets of data were collected for both cautious shifts and risky shifts.

SOCIAL COMPARISON

Stoner's subjects were graduate students in the School of Industrial Management at MIT, and when members of the school first heard about the outcome of Stoner's experiment (then thought of as risky shift), they argued that it could be explained by the fact that the field of industrial management sets a positive value on the ability to take risks. It is a part of the role of an industrial management student to favor risky decisions, they held. Of course the risky shift was soon produced with liberal arts students, both male and female, and by 1982 had been produced with many kinds of persons in many countries, so nothing specific to students of management could possibly matter. If a value on taking risks were involved, it would have to be a very widely held value. Such a value might exist, but even if it did, how would it help to explain the shift that is the phenomenon? If someone thinks risk-taking is admirable, he thinks it when he is alone making the initial individual decision just as much as when he assents to the group consensus or, also again, makes the final postdiscussion decision. That line of thought looked profitless.

An accidental observation of an atypical group suggested a way in which a value on risk might cause a risky shift to be produced by discussion. Group dis-

cussions of the Choice Dilemmas practically always begin with an "initial ballot" in which each participant states the odds he favors. On one occasion a group I was observing neglected to do that. One member of the group made extravagant use in the discussion of what can be called the "rhetoric of risk." "Life is not static; it's dynamic. You have to take chances. Nothing ventured, nothing gained!" When the group was ready to try for a consensual decision, they finally thought to take a poll of initial opinions. The rhetorician turned out to have selected odds of 7 in 10, the second most conservative position in the group. He had clearly conceived of his position as daring, but until all decisions were made public he had no way of knowing whether 7 in 10 was risky or cautious for the problem being discussed.

What if everyone wanted to take a moderately risky position on Stoner's Choice Dilemmas, and what if everyone thought he *was* taking such a position when he wrote down his initial private preference? It would not be possible for everyone in the group to be on the risky side of the mean of the preferences of the group. Really, no one could tell what a risky position would be on novel problems before the distribution became known. Once the real location of the mean was known, should it not be the case, granting that everyone wanted to see himself as reasonably audacious, that those who were really below the mean would be motivated to adopt riskier positions and so change the mean and produce the risky shift? It would all be a process of self-presentation and social comparison (Festinger, 1954), and the effective function of the group discussion would simply be to teach each person how to present himself (to himself and to others) in the way he wished.

The paragraph above states the kernel of the social comparison theory, but it was formulated before all the facts were known. Shifts were as likely to be cautious as risky, so no general value on risk-taking could explain all shifts *in* risk. But the idea of an unqualified value on risk-taking was always silly—one sees in retrospect. No one wants to see himself or have others see him as a big risk-taker when a life is in danger and the welfare of dependent children and elderly parents is part of the stake. In such circumstances caution is a virtue. In fact, a social comparison theory adequate to the phenomenon of shifts in risk as we now know it must hold that the value engaged, the relevant virtue, is a function of the particular choice dilemma. For some problems one wants to be like everybody else, to conform to the central tendency; for others one wants to be on the risky side of the central tendency, though not so far out as to seem foolhardy; on some problems one wants to be on the cautious side, though not so far out as to seem cowardly.

The theory of social comparison can be stated in three interlocked propositions.

1. Any given choice dilemma evokes in most people the desire to be risky or cautious or average.
2. In advance of information as to the distribution of positions on the

choice dilemma, each person imagines that he does, in fact, occupy the kind of position he considers desirable.
3. When the actual distribution of positions on the choice dilemma of all individuals in a group becomes known, those who are not where they want to be, and thought they were, will be motivated to change, thereby producing group mean decisions that represent, according to the individual dilemma, a risky shift, a cautious shift, or no shift at all.

What kinds of evidence support the theory? One claim the theory makes is that people think a more-than-average ability to take risks is an admirable quality for those problems on which a risky shift occurs. Why should they do so in, for instance, the case of Mr. A, the electrical engineer? Whether or not he can afford to take the chance of going with a new company that has a low probability of success but promises a terrific payoff if it does succeed, should, by anyone's reasoning, depend on how good an electrical engineer Mr. A is. If he is exceptionally intelligent, well-trained, and creative, he really need not worry about the company's chancy future, because there will always be a good position for a really top engineer. If his competence is marginal and he considers himself lucky to have the job he has, then he probably would be wiser to hang onto it and move only if a very sure thing is offered. The experimental subject in the role of adviser to Mr. A has to imagine A's competence, and that he does by projecting his own, as he sees it, and so a recommendation of high risk will suggest high competence. That line of thought has been very directly tested in several clever ways (Jellison and Riskind, 1970, 1971; Jellison, Riskind, and Broll, 1972).

Using just those ten of the twelve original Choice Dilemmas which reliably produced risky shifts, subjects read through booklets purportedly filled out by individuals of whom the subjects were to form general impressions. One set of booklets contained ten very low-risk answers (a mean of 8.0), and one contained very high-risk answers (a mean of 2.4). After reading his booklet, each subject was given a rating form on which to record his impression of the person who had filled out the booklet. The form included the adjectives "clever," "creative," "ingenious," "innovative," "intelligent." The high risk-taker was rated as significantly higher on all those aspects of ability than was the low risk-taker. The ten items of the original twelve are all items on which individual ability ought to affect willingness to take risks, and high risk-takers were, accordingly, credited with higher ability than low risk-takers.

The social comparison theory holds not only that risk-taking will be admired with respect to items that show risky shifts, but also that caution will be admired on items that show cautious shifts. That claim has been tested by having subjects first answer a set of choice dilemmas (both risky shift and cautious shift dilemmas) and then go through the same dilemmas answering in the way "they would most admire." Many laboratories have carried out this exercise with largely uniform results (Levinger and Schneider, 1969; Myers, 1982).

For items that reliably produce risky shifts (e.g. in Stoner's Choice Dilemmas, all but No. 5 and No. 12), the most admired answer is consistently riskier than a subject's own answer, and that result strengthens the conclusion that risk is valued on those dilemmas. For items that produce cautious shifts (No. 5 and No. 12 in Stoner's set and numerous dilemmas written by others) the most admired answer is often more cautious than the subject's own answer, but it is sometimes simply not significantly different. The evidence that caution is valued on cautious shift items is not so strong as the evidence that risk is valued on risky shift items, but that fact actually supports the social comparison theory. It does so because one finds in the literature (e.g. Fraser, 1971; Fraser, Gouge, and Billig, 1971) reports that cautious shifts are not so large as risky shifts and not so regularly obtained, on the items where they should be obtained, and Fraser has speculated that risk has a certain attractive daring quality in some cultures even when it entails irresponsibility.

Another claim of the social comparison theory is that most people think they are superior to the average in ability to take risks when risks should be taken and in ability to be cautious when one should be cautious. The claim is tested by asking people first to answer a set of choice dilemmas (both risky and cautious) and then to go through a second time, answering as they think the average person would do. The test has been made many times (e.g. Hinds, 1962; Levinger and Schneider, 1969; Myers, 1982; Wallach and Wing, 1968), and the claim seems always to be confirmed; in advance of discussion almost everyone thinks his decision better realizes the ideal the item engages than does the answer of the average person. Since not everybody can be better than the average person, there is something to be learned when positions are made known.

The credibility of social comparison theory is strengthened by the fact that the tendency to think oneself somewhat better than the average of a reference group is extremely general. Jean-Paul Codol (1975) of the Université de Provence has provided a massive demonstration (some twenty studies) of the phenomenon he calls the "superior conformity of the self" or the "PIP effect" (for *primus inter pares* or "first among equals"). Codol's view is that with respect to behavior or traits expected of a reference group and valued by the group, each person tends to see himself as like the others, only a bit better. Codol has demonstrated the PIP effect in children at various grade levels; in adults who are legal, medical, and academic professionals; and in trade unionists. He has, in addition, put together groups of strangers and created a premium on accuracy or on cooperativeness or on competitiveness, and, sure enough, each individual saw himself as closer than the others to the ideal created.

The final claim of social comparison theory is that when the actual distribution of decisions is made known in group discussion, it is that information which serves to produce a shift either to increased risk or to increased caution. It was Teger and Pruitt (1967) who first thought of the best way to test the claim. If it is the information on others' positions that causes shifts, then that

information alone, with no discussion whatever, should be sufficient to produce the effects. With all discussion forbidden, subjects simply held up cards on which appeared the odds each favored. That position information alone was sufficient to produce significant risky shifts on the usual ten Stoner problems and also to produce cautious shifts (just short of significance) for No. 5 and No. 12. The size of the shifts was, however, only half that produced by discussion, so it is clear that position information is not the only important product of free discussion.

The Teger and Pruitt findings have been very often replicated (e.g. Myers and Bishop, 1971; Myers, 1978; Stokes, 1971), but there have also been attempts at replication that failed (Pruitt, 1971a, 1971b). Most of the failures occurred in the 1960s and early 1970s, prior to the discovery of two factors that may have caused the social comparison effect to appear to be less robust than it is. It occurred to Myers that the initial decision a subject makes, before he learns the decision of others (in the standard experimental design), may function as a kind of commitment causing the individual to resist change, especially as his only reason to change would seem to be information about what others think. Therefore Myers, Bach, and Schreiber (1974) tried a different experimental design (called a "between-groups" design) in which all subjects simply made decisions once, but some (randomly selected) did so with information about the decisions of others and some without such information. With that design, the ability of positional information to produce shifts was greater than usual, and Myers (1982) thinks it is clear that the initial commitment decision dampens the social comparison effect.

PERSUASIVE ARGUMENTS

Suppose a problem has an objectively correct answer that is not known to all the subjects available for group discussions but is known to some of them. Here is such a question: "If you were to go due south from Detroit, Michigan, what would be the first foreign country you would reach? Brazil? Cuba? Mexico?" The answer is Canada; the city of Windsor, Ontario, because of a twist in the United States-Canadian border, lies south of Detroit. If each member of a population of subjects answered independently, then those who knew the correct answer would give it and the rest would not. If the subjects were then convened as small groups to discuss the question, there would be some groups in which one or two members knew the answer and the rest did not. The correct answer would prevail more often than the various incorrect answers. This answer—Canada—would usually not represent the most common prior response of the members, and so its victory could not be attributed to majority pressures. Neither is "Canada" the mean or central tendency of the initial opinions, and so the simple impulse to converge will not account for the agreement on "Canada." "Canada" is a minority answer, on one side of the mean, which happens to be correct. If the decision to take a high risk on the Stoner problems

could be considered objectively correct, then the shift to risk would be explained.

Surely it is pointless to compare choice dilemmas to questions of fact concerning geography, history, politics, and the like, because choice dilemmas have no right answers in the sense that the question about Detroit has a right answer. Notice, however, that a group of persons asked to identify the first foreign country south of Detroit would have no way, while in discussion, of checking on the correctness of the unexpected suggestion: "Canada." Probably, however, all would be persuaded to adopt that answer. Why should they be? It is likely that the person who made the suggestion, knowing that it can be checked on a map, would make his suggestion with great confidence, and confidence is a powerful persuader. The Canada answer would also be persuasive by virtue of its paradoxical quality. Since Canada, in general, lies north of the United States, and everyone knows that, anyone who asserts that a part of Canada lies south of a part of the United States must be presumed to have good reason for his assertion. Evidently information can be persuasive even when it is not known to be factual.

Look once again at the sample discussion of the problem of Mr. A, the electrical engineer, attending now to the content. Participant F says: "I read in *The New York Times* or somewhere that there is a big demand for electrical engineers. Even if this company should fall through, he can always get a job, probably at the same salary. I figure even if the company goes broke, he can still claim valuable experience in looking for another job later on." Nothing in this contribution identifies the correct answer to the problem of Mr. A, the minimal odds that he should accept. However, if you had not thought of F's argument—that engineers are in short supply and can easily find jobs as good as the one A now has—then this argument should operate as a *correction* on whatever odds you initially chose. The argument exercises a force toward greater risk, and if it has not figured in your thinking and you find it persuasive (citing *The New York Times* would help make it so), then the argument should move you toward the acceptance of lower odds, whatever your initial thought.

That much of the persuasive arguments theory was worked out when the phenomenon to be explained was still the risky shift. Group discussions of the dilemmas were not altogether different from discussions aimed at solving problems of fact or logic, because while it was true that the dilemmas had no objective correct answers, there were always relevant arguments that were more or less persuasive. It seemed reasonable to suppose that each individual prior to discussion would think of some relevant arguments, but not all, and that the arguments one thought of would not be identical with those another thought of. In group discussion each would tell all he knew, and that would shift group opinion.

From these beginnings a really beautiful set of ideas dawned (Bishop and Myers, 1974; Vinokur and Burnstein, 1974). Decision shifts, we knew, could be either to risk or to caution and could be large or small, depending on the item.

Suppose any given item could be said to have a pool of possible arguments, a latent population of relevant arguments, a below-ground bulb structure of relevant arguments. And suppose the total pool for a given item favored risk or caution in a given degree. A tendentious pool of that kind would be analogous to the correct answer that an objective question has. Suppose, further, that each individual who ruminates on an item draws a sample of relevant arguments from that item's pool. Insofar as the pool favors risk, weakly or strongly, his sample will do the same, and insofar as the pool favors caution, weakly or strongly, so will each sample. The samples drawn from the pool will determine the initial individual decisions, and so the mean of these should be more or less risky or cautious according to the character of the total pool.

Assuming that individual samples of arguments are not identical, that arguments are only partially shared, in group discussion all samples will be combined and all samples together will better represent the character of the total pool than individual samples. The mean of individual initial decisions correlates highly with the size of the shift after discussion simply because extreme initial means must be drawn from extreme pools and moderate initial means from moderate pools. And so discussion will shift the mean of individual initial positions farther out in the direction of initial inclination.

1. For each choice dilemma there exists a pool of arguments favoring risk or caution, and the number and persuasiveness of the arguments varies from item to item. The character of the pool for a given item must be presumed to be culturally and historically relative even though for twenty years in many countries there is little evidence of variation.
2. Each person who considers a given dilemma thinks of some sample of the arguments in the total pool, but not all. The arguments he thinks of determine his initial individual decision. The arguments considered by any set of individuals will be only partially shared prior to discussion.
3. The important thing that happens in discussion is that individual arguments are expressed and become fully shared. Because the choice dilemmas have total argument pools in which the balance favors either risk or caution, the larger sample of the pool made available to all in discussion will produce either a risky shift or a cautious shift according to the direction of prediscussion inclinations as revealed in the means of initial decisions.

What sorts of evidence support the persuasive arguments theory? The theory claims that each Choice Dilemma is associated with a distinctive pool of relevant arguments, which may favor risk or caution strongly or moderately, and that the means of the initial decisions (because they are based on samples from the pool) should reflect the character of the pool. Vinokur and Burnstein (1974) asked subjects first to respond individually to each of five dilemmas (two risky, two cautious, and one neutral) in the usual way. Subjects were then asked, without benefit of any discussion, to list all the arguments they could

think of relevant to each item, indicating whether an argument favored ris[k] or caution. Those combined complete argument lists can be thought of as a[n] approximation to the total pool characteristic of each item. Choice dilemma[s] known to shift to risk had item pools in which the proportion of risky argu[ments] was greater than it was in the pools for cautious shift items, and the pool[s] for the neutral items was almost exactly evenly balanced between risky and cautious arguments. In addition, the means of the individual decisions on the[se] items were correlated with the proportion of risky to cautious arguments in the item pools.

The persuasive arguments theory further predicts that the persuasive argu-ments constituting the pool for a given item will be only partially shared among persons prior to group discussion of the item. Vinokur and Burnstein had judges classify together arguments that were essentially identical and so obtained lists for each item of distinct arguments. They were then able to deter-mine how widely each argument was shared. If all arguments for any given item were known to all those convened for a discussion, no sharing would be created by the discussion and so no shift should occur. In fact, sharing was only partial for each item; indeed, no argument was shared by more than one-third of the subjects. So partial sharing in advance of discussion, which is a *sine qua non* of discussion-induced shifts, was, in fact, the case.

Finally, the ratio of risky to cautious arguments in an item's pool should predict the size of shift produced in discussion, since each participant will increase his sample from the total pool, and that sample will be biased toward risk or caution as the pool is. This discussion necessarily had to involve subjects other than those who produced the complete lists of arguments, as the listing process would be expected to make shifts unlikely. The discussions of the items were created by Vinokur (1971) for another experiment, and the mean shifts for each item were taken from that experiment. The size of the shift on an item was predicted by the proportion of risky to cautious arguments in its pool—exactly as the theory requires.

The acid test of the persuasive arguments explanation parallels that for social comparison: Is the one critical variable sufficient in the absence of the other? Are arguments alone sufficient to produce decision shifts? It has been repeatedly and consistently shown that they are. Clark, Crockett, and Archer (1971) had groups discuss choice dilemmas (a subset of six) with the restriction that no one was on any account to say what recommendation he favored. That arguments-only discussion produced as large a risky shift as did full discussion with individuals stating the odds they favored. Myers, Bach, and Schreiber (1974) and Burnstein and Vinokur (1973) have shown in several different clever ways that exposure to arguments alone is enough to produce significant shifts.

One study (Ebbesen and Bowers, 1974) stands out for me, because it is espe-cially elegant and also because it will help us to explain *Twelve Angry Men*. This experiment pulls apart the natural risk position of a dilemma, as determined by its pool of relevant arguments and the ratio of risky to cautious arguments actu-

generated in a discussion. Five of Stoner's problems were used, chosen so [as] to sample the range from high risky shift to high cautious shift. For each of the items the experimenters created five different scripted discussions, con-[tr]ived so as to vary the risky-to-cautious argument proportion. In each scripted, [re]hearsed, and tape-recorded discussion just ten relevant arguments were used. [F]or each item there was a discussion in which the risky-to-cautious proportion was .90, which meant nine arguments favoring risk and one favoring caution. For the same items there were discussions in which the proportions were: .70, .50, .30, and .10. Discussions with arguments in just the same five proportions, then, were created for each item in spite of the fact that the items had quite dif-ferent natural "bents." Subjects were passive auditors who first answered an item for themselves, then listened to one of the five discussions of that item, and finally answered again for themselves. The question was whether the con-trived argument proportions would predict the direction and size of the shift. Figure 6–3 shows how well they did so.
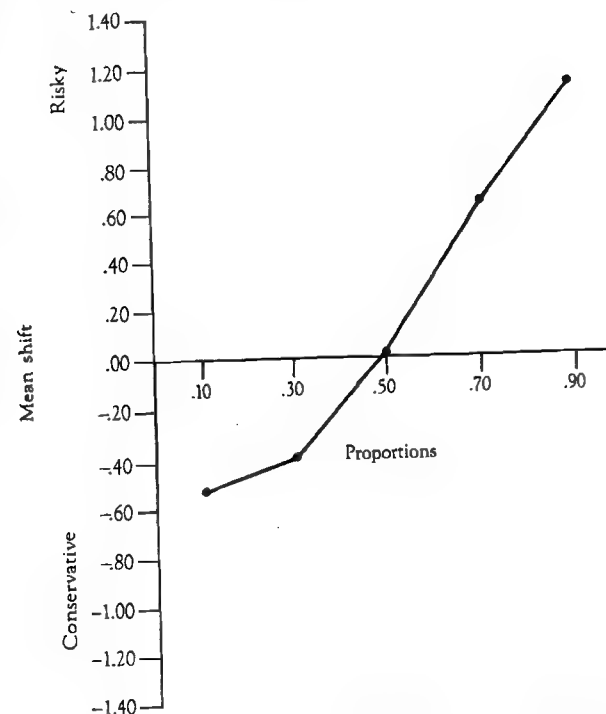

Figure 6–3. Average shift in risk estimates as a function of the proportion of risky arguments that subjects heard

When half the arguments favored risk and half caution, the shift was exactly .00. As the proportion of arguments favoring risk increased, so did the shift to increased risk, and as the proportion of arguments favoring caution ("conservative") increased, so did the size of the shift to increased caution. As you can see in Figure 6–3, however, shifts to caution were not as large as shifts to risk. I take that to be yet another manifestation—and there are many running through the study of group polarization—that the value on risk is in some general way more powerful than the value on caution.

Decision shifts were produced by arguments actually heard and not by the natural bent of an item, not by the total pool of arguments associated with an item. That is very interesting, because it shows that if we think of the pool of relevant arguments as something like the correct answer for that item, it is possible for an odd, unrepresentative discussion to produce an incorrect answer. I shall want to argue that precisely this must be supposed to have happened in the trial that precedes the jury deliberations in *Twelve Angry Men*. We have to suppose that, because Henry Fonda, the lone juror voting "not guilty," is in the course of the play able to produce powerful evidence and arguments, not brought out in the trial, in favor of that verdict, arguments so powerful that the audience is convinced that the correct verdict is "not guilty." For some reason the pool of arguments that defines the right answer for the young defendant accused of murder was not properly sampled in the trial itself. There are several remarks in the play to the effect that the defense did a poor job, which in our terms must mean a failure to penetrate to the pool of evidence and arguments that establishes "not guilty" as the correct verdict. The inadequate defense leaves the way open for Henry Fonda to re-try the case and come up with the right answer.

Polarization in General

It is odd really that most of the work using choice dilemmas (Stoner's or others) has also presented subjects with choice options in the awkward form of odds ranging from 1 in 10 to 10 in 10. The form is awkward because low odds represent high risk, and so an unnecessary though simple mental transformation must be made. More important, there is no printed option that is clearly intended to represent psychological neutrality. For those reasons I am happy to introduce work done at the University of Bristol by Colin Fraser and his associates (Fraser, Gouge, and Billig, 1971). They used eight dilemmas, four risky and four cautious, some of them Stoner problems and some new problems written by the authors. They first used the familiar odds options (1 in 10, 2 in 10, etc.) and replicated all the usual shifts and correlations. With new subjects they used the same problems, but response options were a simple seven-point scale. For Mr. A, the electrical engineer, the riskiest choice was "1. Strongly recommend Mr. A take the new job," and the most cautious was "7. Strongly recommend Mr. A remain in present job." Positions 2 and 3 were labeled so as to represent, respectively, moderate and little risk, and positions 5 and 6, little and moderate caution. Position 4 was labeled "Neutral, i.e., the two alternatives appear equally balanced."

The seven-point scale with a labeled position of neutrality produced results exactly comparable with those produced using the odds alternatives. It was easier to answer, and the neutral point on the scale functioned as the psychologically neutral position: Dilemmas that scored between 3.5 and 4.5 on the initial individual means showed no significant shift as a result of discussion. The odds options and the seven-point scale were proved to be essentially equivalent instruments for the study of problems involving risk. However, the seven-point scale has a wider range of usefulness not shared by the odds options. A seven-point scale of the kind Fraser used is called a Likert-type scale (named for Rensis Likert, a pioneer in opinion survey work), usually used to measure attitudes, not risk options. The use of the seven-point scale clears the way for a new reformulation of our decision choice phenomenon, a reformulation having nothing to do with risk and, indeed, not linked to any one kind of content. A seven-point scale can even be used with the meaning: "1. Very definitely guilty"; "4. Equally likely to be guilty or not guilty"; "7. Very definitely not guilty." We shall shortly see it used in that way and so draw near to *Twelve Angry Men*.

It was Moscovici and Zavalloni (1969) of the Laboratoire de Psychologie Sociale in Paris who first suggested that the much-studied shifts in risk might be an instance of a completely general group polarization effect. Perhaps it is the case with every sort of attitude, belief, and decision that group discussion moves the mean of initial individual decisions farther out, away from the mean, in the direction of the initial inclination. The Paris group used Likert scales running from +3 to −3, rather than from 1 to 7, which was Fraser's choice; it makes no difference, and since most work since has adopted the Paris form, we shall do the same. A scale item reads like this:

- − 3: strongly disagree
- − 2: disagree
- − 1: slightly disagree
-   0: incapable of expressing an opinion
- + 1: slightly agree
- + 2: agree
- + 3: strongly agree

With items in this form the group polarization hypothesis can be very simply stated: On whichever side of zero the initial mean falls, it will, after discussion, move farther away from the mean toward the nearer pole.

The first experiments were done with secondary school students (eighteen or nineteen years old) in Paris using discussion groups of four members each. Attitudes were initially assessed toward General de Gaulle and toward "the

Americans." A typical item on de Gaulle reads: "He is too old for his important political task" (agree or disagree on a scale from $+3$ to -3). A typical item on "les Américains": "American economic aid is always used for political purposes." The average initial attitude toward de Gaulle was mildly favorable ($+.90$); after discussion both the group consensus and the mean of final individual decisions became more favorable ($+1.18$), which is to say that discussion moved opinion farther in the direction of the nearer pole. Initial attitudes toward Americans were mildly negative ($-.60$) at first and became more so (-1.09) after discussion, which is to say that they moved farther out from zero in the direction of the original departure from zero. While we are in Paris, we should add the study by Doise (1969), which showed group polarization in the case of attitudes of students at a school of architecture toward their school; initially negative, they became more so after discussion.

The work we have considered until this point all concerns problems or objects of attitudes that elicit the same initial inclination from all the groups studied, and that leaves us wondering how group polarization would work with social issues that arouse a range of opinion, including opposed polarizations. For instance: racial prejudices. Myers and Bishop (1971) had high school students respond to a 100-item prejudice inventory, and they then created two sorts of homogeneous populations: the more prejudiced and the less prejudiced. The mean scores of the two populations fell on either side of zero. What would happen to individual attitudes if like-minded persons were formed into small groups and given some statements relevant to prejudice to discuss? The theory of group polarization holds that because of social comparison and the sharing of relevant arguments, the initially divergent groups should move toward their individual poles and, therefore, end up farther apart than they were before discussion.

The groups were given eight propositions to discuss, for instance: "Some people recently have been saying that 'white racism' is basically responsible for conditions in which Negroes live in American cities. Do you agree or disagree?" However, everyone was first asked to respond for himself to each item on a Likert-type scale. After discussion, they were asked to respond again. Both homogeneous groups moved toward their respective extreme poles, and so they were farther apart after discussion than before. Notice that these discussions did not put the prejudiced and unprejudiced together, the kind likely to take place when, for instance, a neighborhood action related to prejudice is to be voted on. The discussions in the experiment were the kind that would take place in advance of any official representative meeting, discussions in which, for the most part, like-minded individuals get together and exchange views. It is perhaps the first principle of social psychology that people associate with people who are similar to themselves, and so most informal discussion of issues goes on in homogeneous groups. Such discussions should, the theory holds, *amplify* initial disagreements between opposed groups.

When we are interested in "cooling off" an issue, in depolarizing extreme factions, and that is certainly something we are often interested in doing (the middle East, relations between the United States and the U.S.S.R., prison riots, Italians and blacks in Boston), the cry that always goes up is "keep them talking." Evidently the layman's theory is that if extreme factions are brought to talk together, there will be a net depolarization, a movement away from both extremes toward a moderate position. Eugene Burnstein (1982) of the Research Center for Group Dynamics at the University of Michigan has recently reported some results on depolarization that are so suggestive they are sure to be followed up in many experiments.

Burnstein has been the principal champion of the persuasive arguments explanation of shifts in risk, and his studies on depolarization were done with four standard risky shift choice dilemmas, one standard cautious shift dilemma, and two standard neutral (or no shift) dilemmas. He created six-member groups to discuss the items, but those groups, unlike those put together to demonstrate group polarization, were not homogeneous. They were made up of two sets of opposed extremes (three members each). What happened? The principal effect was depolarization or convergence to the central tendency, and it was a big effect—two or more units on the odds scale. So the layman's prescription for damping contention seems to be right. However, even the first experiments indicate that two qualifications must be entered.

Neutral items are items that have argument pools evenly balanced between risk and caution. When polarized factions are put together, the only thing that happens on those items is depolarization or convergence. Risky items are items having pools in which arguments for risk are more numerous and persuasive than arguments for caution, and for cautious items the balance of arguments in the total pools is just the reverse. In a certain sense, remember, risky and cautious dilemmas can be said to have "correct answers"; the correct answer is just the total pool of relevant persuasive arguments. What happens when problems like these are discussed by opposed extremes with equal numbers representing each extreme? Depolarization is still the largest effect—the extremes move toward one another—but it is not the only significant effect. The naturally risky items also all show a risky shift and the naturally cautious items a cautious shift. This is to say that the new, converged-upon moderate position is for the risky items closer to the initial position of the risky extreme than to the initial position of the cautious extreme, and for the cautious items it is nearer the initial cautious extreme.

Here is an example: On dilemma A, a risky shift dilemma, the two extremes were 5.09 points apart initially and only 2.00 points apart after discussion. They closed 3.09 points of difference. However, they did not evenly split the difference. The cautious faction moved more than the risky faction, with the result that the final mean position (3.50) was riskier than the initial mean position by 1.21 points. For cautious item E it was the risky factor that did the

greater part of the moving. What that means is that the final outcome is a kind of compromise between depolarization or convergence and the right answer, so that the final position is more moderate than either extreme but is nearer the extreme that argued for the correct answer or, more exactly, the extreme that had the better pool of arguments to draw from. The action implication would seem to be: Bring extreme factions together if you want a decision more moderate than either extreme, but if one extreme is more nearly correct or has a better stock of arguments, the final common position will suit them better than the opposition.

Of course, depolarization on long-standing real social issues involving race, gender, capital punishment, and the like is not so easily accomplished as it is, with story problems involving risk. Burnstein has identified one reason why not (we know there are many others). He compared the amount of depolarization produced when extremes discussed long-familiar, much-debated issues (including capital punishment) with the amount produced for new problems. On unfamiliar problems massive depolarization occurred, but on familiar social issues, little or none. That tells us what we know; that getting three Israelis and three members of the PLO to talk together will not produce sweet harmony. We knew that already, but persuasive arguments theory tells us one general reason why familiar and long-debated issues do not depolarize easily. It must be, in part, because the total pool of arguments has long been familiar to all. There is nothing new to be learned in group discussion.

Group Polarization in Jury Deliberations

A jury must deal with a choice dilemma for which the response options are verdicts of "guilty" or "not guilty." The problem a jury faces is not a one-paragraph story or a one-sentence statement of opinion. It is nothing less than the trying of a case: evidence, testimony, argument, and rebuttal, extending sometimes over many days. The trial each juror hears disposes him more or less strongly to one or the other of the two verdicts, and each goes into the jury room with an initial prediscussion opinion, which is often promptly disclosed on a first ballot. The discussion that then takes place is designed to produce the level of consensus the judge has instructed them to reach; for six-member juries it must be unanimous, and for twelve-member juries it is often the same, but majority decisions are also sometimes legal. In jury discussions both persuasive arguments and social comparison processes occur. Jury deliberations are therefore the sort of situation in which group polarization should occur, and Myers and Kaplan (1976) have shown that it does—at least in an experiment.

The law requires that jury deliberations be private, and so they may not be directly studied. The research approach, therefore, is to simulate the real thing. Myers and Kaplan did not attempt a very realistic simulation: Juries were ten groups of six undergraduate subjects each, and the cases were printed adapta-

tions taken from the *California Law Reviews*. In Chapter 8 some extremely realistic simulations are described in which jurors were actually drawn from the juror rolls, and cases were carefully reenacted in courtrooms with real judges and attorneys. In order to see whether group polarization occurred in jurylike deliberations, Myers and Kaplan judged that elaborate efforts at realistic simulation were not necessary.

Four traffic felony cases abstracted from the *California Law Reviews* were used. Each case was written in two forms: High incrimination or low incrimination. Here is one case in the high incrimination form; it created a very strong presumption of guilt:

L., a bakery truck driver, was charged with negligent death in the fatal injury of a 2-year-old child, C. The defendant was engaged in driving a bakery truck which made both home deliveries and "on the street" sales. Purchasers "on the street" were generally children. The truck had made regular visits to this particular street and was well known to local children. On the day in question, L. parked parallel to the curb. No other vehicles were parked in the vicinity of the truck. L. noted a group of children playing in a nearby backyard and rang some chimes, the customary means of attracting customers. The children approached the truck, on the street side, and made several purchases. While selling the children the bakery products the driver's back was turned to the front of the truck. In this period, C., who had been playing with the others, approached the truck and began playing in the street, in front of the left front wheel.

Due to his position while serving the children, and the size of the truck, it was impossible for the defendant to see the child, unless he walked around to the front.

After serving, the defendant returned to his seat in the truck. The truck was deemed by safety officials to have adequate windshield area and mirrors to assure good visibility on all sides. The driver looked into the side mirror to ascertain that the children he had served were standing on the sidewalk, but in testimony, couldn't remember whether he had looked in front of the truck for children. C. was 3 feet tall, and would have required a downward glance from the front seat to be seen.

Since children were the primary customers of the truck, police officers had instructed the driver on several occasions to blow his horn when leaving a parking space. Adult witnesses testified that the defendant did not, on this occasion, blow his horn. The driver started the truck, checked his rear view mirror, and pulled out into the street. While pulling out, his head was turned to the rear, watching for traffic. He stopped when he heard the child's screams. Front and rear wheels had passed over the child. [Kaplan and Kemmerick, 1974]

In the low-incrimination version of this case the driver blew his horn several times, checked his rear view mirror, and also looked through his windshield in front of the truck but failed to allow for a blind spot just forward of the front wheels that resulted from a peculiar driving position. "Jurors" read one or the other version of a case and then rated the degree of guilt of driver "L." from "0 (Definitely not guilty)" to "20 (Definitely guilty)." In addition, they were asked to assume that L. had been found guilty and to recommend a level

of punishment from the minimum the law allowed for the charge (1) to the maximum (7). Those judgments, made for both versions of four cases, were the individual prediscussion opinions. The cases were then discussed for a few minutes by a "jury," and jurors individually made the two sets of ratings a second time. A control always used in group polarization studies was used also here: Some subjects simply rated the cases twice without intervening discussion so that any possible effect of familiarization with the materials could be checked.

Initial mean guilt ratings across high incrimination (or high-guilt) cases were about 12.5 (on the "Definitely guilty" side of the midpoint 10), and means after discussion moved toward the extreme to a mean value of about 15.0. Low-guilt cases were eventually rated at about 6.0 (on the "Definitely not guilty" side of the midpoint) and after discussion had moved out to about 4.0. In short, a polarization effect of substantial size occurred, with the "initial ballots" turning into group verdicts more extreme in the original directions. Comparable polarization effects occurred for the levels of punishment recommended, and the controls for familiarization showed no significant effects at all. All results are pictured in Figure 6-4.
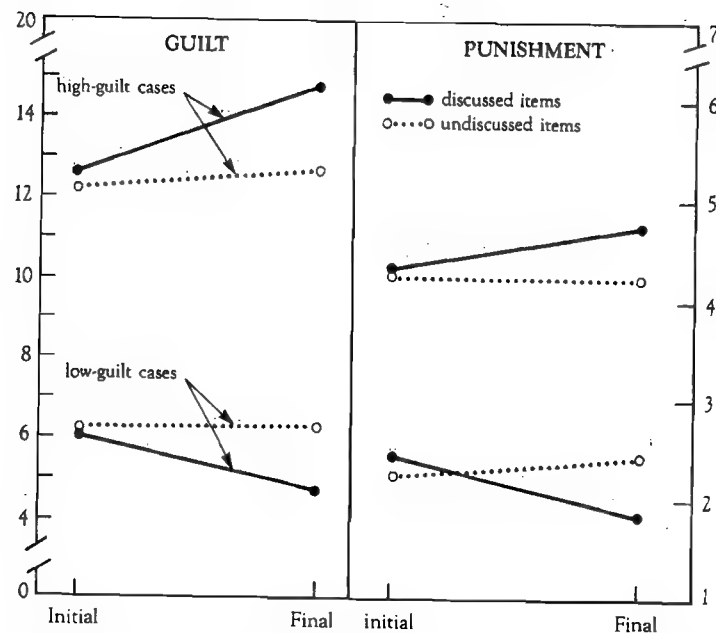

Figure 6-4. Initial and final ratings for discussed and undiscussed cases

(D. G. Myers and M. F. Kaplan, Group-induced polarization in simulated juries, *Personality and Social Psychology Bulletin,* 2 [1976], pp. 63–66. Copyright © 1976 by the Society for Personality and Social Psychology, Inc. Reprinted by permission of Sage Publications, Inc.)

The Myers and Kaplan study seems to be the only jury simulation study that was set up to test group polarization, but just about all mock jury studies have incidentally collected data that make possible a check on group polarization: individual predeliberation opinions and a postdeliberation group verdict, often also individual postdeliberation opinions. In every instance where the report of data makes it possible to check, group polarization occurs (e.g. Davis et al., 1975; Hastie, Penrod, and Pennington, 1983; Kerr and MacCoun, 1985; Padawer-Singer, Singer, and Singer, 1977; Saks, 1977; Valenti and Downing, 1975).

Of course, even the most realistic simulation is not a real jury-of-record, making a decision with consequences for a real defendant, but we can be sure that group polarization occurs also in real juries because of a very telling statistic on their outcome. While it is unlawful to "tamper" with real juries by studying them when they are doing their work, it is entirely lawful to question jurors after they have discharged their obligations, and that was done on a massive scale by Kalven and Zeisel (1966). The vital statistic is that for 90 percent of juries that must reach unanimous agreement and do not hang, the final verdict is consistent in direction with the majority on the initial ballot. It is very rare for an initial ballot to yield a unanimous outcome; initial ballots are less extreme than that. Therefore, to know that the predeliberation majority, whether for conviction or for acquittal, predicts the final verdict 90 percent of the time is powerful presumptive evidence that group polarization occurs in real juries.

Twelve Angry Men

The vital statistic from Kalven and Zeisel that 90 percent of real juries finally vote in the direction of their initial majorities restores in full the apparent absurdity of the vote switch in Reginald Rose's play. So, apparently, does the entire story of group polarization, since Rose's twelve-man jury precisely *does not polarize.* It does not become more extreme in the direction of original inclination but reverses an original inclination that was initially almost unanimous. What then can be the purpose of linking a silly piece of fiction with a tradition of good cumulative research, and how is it possible that the research should prove the verisimilitude of *Twelve Angry Men?* The congruence is on the level not of surface outcomes but of causal processes. Persuasive arguments and social comparison operate in the play in such a way as to make the vote switch the correct outcome. And the play is more than an illustration of the two principles of group dynamics; it expands our understanding of them.

A trial has just ended as the curtain rises; a trial of murder in the first degree, premeditated homicide, the most serious charge in our criminal courts. "If there is reasonable doubt in your minds," the judge instructs the jury, "as to the guilt of the accused—then you must declare him not guilty. If—however—

there is no reasonable doubt, then he must be found guilty" (Rose, 1956, p. 9). The verdict must be unanimous.

What are the principal facts in the case? A man has been stabbed to death, and his nineteen-year-old son is the accused. They are members of some unspecified ethnic minority, living in a poor section of a large city in an apartment that is periodically shaken by the rush past its windows of an elevated train. What pieces of evidence connect the son with the crime? There are several small circumstantial items but only three weighty ones.

The murder weapon is a switchblade knife. Everyone agrees that it is an unusual knife, and a storekeeper has testified that he sold just such a knife to the son, arguably just this knife, it being the only one of its kind he had in stock, and everyone thinks it unique. The son, however, testifies that he bought the knife for a friend and lost it through a hole in his pocket before the murder took place.

In the apartment just beneath the scene of the murder lives an old man who has suffered two strokes and walks, with difficulty, with the help of two canes. The old man has testified that on the night of the murder he heard a quarrel overhead and heard the son scream, "I'm going to kill you," and then a body fell to the floor. He heard someone running down the stairs and, hurrying to the door of his own apartment, looked out and saw the son.

Just at the time the murder occurred, an elevated train thundered past the windows of the father's apartment. In the apartment across the tracks a woman lay in bed unable to sleep, her head beside a window looking directly toward the scene of the murder with the passing train between. At the trial she testified that she saw the murder and identified the son as murderer. The prosecutor demonstrated to the jury that it was, in fact, possible to look through the windows of a moving train and identify someone on the opposite side. That eyewitness testimony weighed very heavily with the jurors who voted "guilty."

Persuasive Arguments

The play does not ask us to accept on faith the credibility of the vote switch that occurs between the initial ballot and the final verdict; that is the event the playwright undertakes to make credible. However, the play does require us to accept something quite improbable, to "suspend disbelief" on one point, which is, in fact, the point that makes it possible for persuasive arguments to operate in the jury room in such a way as to make "not guilty" the correct verdict. We must believe that even though the trial has been long and complex, the defense attorney has done a very poor job; only if a poor job was done in the trial would it be possible for the holdout juror, Henry Fonda, to come up with new evidence and new arguments so persuasive as to convince both jurors and audience. There is passing reference to that improbably bad defense:

JUROR NO. 8: I had a peculiar feeling about this trial. Somehow I felt that the defense counsel never really conducted a thorough cross-examination. Too many questions were left unasked.

JUROR NO. 4: While it doesn't change my opinion about the guilt of the kid, still, I agree with you that the defense counsel was bad. [p. 22]

From the point of view of persuasive arguments theory, the case on trial is a problem for which the total pool of arguments and evidence favors "not guilty." That verdict is, in the persuasive arguments sense, the "correct answer." The trial that has taken place before the play begins should be regarded as strictly analogous to one of the contrived discussions, using a total of ten arguments, divided one way or another between risky and cautious, used by Ebbesen and Bowers (1974) with five problems, some of which had correct answers that were risky and some of which had correct answers that were cautious. The trial that took place is like a discussion contrived to go against the natural bent of a problem. The initial ballot is congruent with the balance of arguments heard, as in the Ebbesen and Bowers experiment, but the way is open for a new trial to be carried on in the jury room, and that trial unfolds in a way that is consistent with the total pool of arguments. In effect, the play asks us to entertain as a premise the possibility of a seriously unrepresentative trial so that the playwright can show us what interesting things follow from such a premise.

The holdout juror (Henry Fonda) systematically destroys the three principal pieces of evidence, in the order of their listing here, which puts the most important piece, the eyewitness testimony, in final position. This is, after all, a play, and a play should build to a climax. If it were within the province of this book to include an explicit mathematical model of the persuasive arguments theory, it would be possible to describe more exactly what Fonda accomplishes. In terms of an information integration analysis (Anderson and Graesser, 1976), each piece of evidence would have both a value (for guilty or not guilty) and a weight. The three pieces of evidence retain their value (if accepted, they argue guilty), but Fonda systematically reduces the weight of each to zero.

First—the unique knife. Fonda (Juror No. 8) asks that the guard bring in the knife so that it may be examined again.

JUROR NO. 4 (holding up knife): Everyone connected with the case identified this knife. Now are you trying to tell me that someone picked it up off the street and went up to the boy's house and stabbed his father with it just to be amusing?

JUROR NO. 8: No. I'm saying that it's possible that the boy lost the knife, and that someone else stabbed his father with a similar knife. It's possible. (Juror No. 4 flips knife open and jams it into wall just downstage of door L.)

JUROR NO. 4 (standing back to allow others to see): Take a look at that knife. It's a very strange knife. I've never seen one like it before in my life.

Neither had the storekeeper who sold it to him. (*Juror No. 8 reaches casually into his pocket and withdraws an object. No one notices him. He stands up.*) Aren't you trying to make us accept a pretty incredible coincidence?

JUROR NO. 8 (*moving toward Juror No. 4*): I'm not trying to make anyone accept it. I'm just saying it's possible.

JUROR NO. 3 (*rising, shouting*): And I'm saying it's not possible!" (*Juror No. 8 swiftly flicks open blade of a switchknife, jams it into wall next to first knife, and steps back. They are exactly alike. There are several gasps and EVERYONE stares at knife. There is a long silence.*) [pp. 23–24]

It turns out that No. 8 found an exact replica of the murder weapon in a junk shop just around the corner. So much for the unique knife. The weight of the evidence connecting the murder weapon to the knife the son bought is now zero. Of course, it is reduced to zero in a dramatic way. I have not said that there was nothing to Rose's play except persuasive arguments and social comparison. He uses them (intuitively), but he is also a dramatist.

The second item of evidence is the testimony of the old man downstairs. Whether or not he could have heard the son scream, "I'm going to kill you" with a train going past, and whether such a statement should be taken as a statement of serious intent in any case, is left moot. Much more serious is his testimony that he identified the son running down the stairs. He said he saw him just fifteen seconds after the body hit the floor overhead. The question is raised by the holdout juror whether an old man who has had two strokes and who walks with two canes could in fifteen seconds rise from his bed, take hold of his canes, and walk 12 feet to his bedroom door and 43 feet more down a hallway to his apartment door. Those numbers let us in for one of those tiresome reenactments favored in courtroom drama, with one juror holding a stopwatch and Henry Fonda acting the old man. His time is thirty-nine seconds, not fifteen.

The reenactment raises reasonable doubt about the old man's testimony. What remains is the testimony of an eyewitness, and eyewitness testimony always counts for more with juries than does circumstantial evidence. A woman across the tracks from the father's apartment has said she saw the son commit the murder through the windows of a passing el train.

Juror No. 4, who is the most intelligent and temperate of those convinced the son is guilty, grants that reasonable doubt has been created about much of the evidence, but he himself remains convinced the verdict should be "guilty" precisely because of the eyewitness testimony. He reviews what the woman said and concludes: "As far as I can see, this is unshakable testimony" (p. 60). At this point—late in Act III—it had begun to look as if a verdict of "not guilty" might be brought in; a majority of nine had come round to that view. However, No. 4's obviously fair and well-reasoned summary, together with his impressive appearance, has shaken all the unsteady voters. There is a long

silence, and then No. 4 suggests that the time may have come to admit that they are a hung jury.

JUROR NO. 2: What time is it?

JUROR NO. 8: Can't you see the clock without your glasses?

JUROR NO. 2 Not clearly.

JUROR NO. 8: Oh.

JUROR NO. 4: Glasses are a nuisance, aren't they?

JUROR NO. 8 (*an edge of excitement in his tone*): Well, what do you all do when you wake up at night and want to know what time it is?

JUROR NO. 2: I put my glasses on and look at the clock.

JUROR NO. 8: (*to Juror No. 2*): Do you wear your glasses to bed?

JUROR NO. 2: Of course not. No one wears eyeglasses to bed.

JUROR NO. 8: The woman who testified that she saw the killing wears glasses. What about her? [p. 61]

The jury takes a few minutes to digest the new information. The woman had worn bifocals in court and had never taken them off. Henry Fonda wraps it all up by saying the eyewitness probably was not wearing her eyeglasses in bed. "Maybe she honestly thought she saw the boy kill his father. I say that she only saw a blur" (p. 62).

In the end, reasonable doubt has been created about every single piece of evidence, and so "not guilty" must be the verdict. If the premise is accepted that the trial that preceded the jury deliberations was unrepresentative of the total pool of arguments, then it is credible that a new sample from that pool, drawn in the jury room, might be more representative and lead to the correct answer. Perhaps it is unrealistic to have one juror think of most of the new arguments, but that one juror is the star of the show, the hero, Henry Fonda. For once the heroism on view is largely cognitive or intellectual in nature. The most extraordinary thing about No. 8 is that he comes from the trial which all twelve have attended with something like a full set of the persuasive arguments that make "not guilty" the correct answer whereas the eleven others do not. However, No. 8 is heroic in character as well as intellect and so a second process of group dynamics operates in the play.

Social Comparison

Social comparison operates in *Twelve Angry Men* in a way that is more subtle and more general than its mode of operation in problems involving risk or in discussions of social issues. Indeed, I think the play reveals better than any experiment the real domain of social comparison in group decisions. Imagine that you are a member of some decision-making body, a board or a committee, and that you are expected to vote on every issue that comes before the group

even though you privately know that you are not always well qualified to do so. That happens. It has happened to me fairly often. One can be elected or appointed to a position on the basis of presumed qualifications and yet not necessarily have all the qualifications. I have voted to accept dissertations when I did not fully understand some technical matter controversial among the readers. I have voted for or against the award of fellowships in areas entirely outside my knowledge: theoretical physics or the pronominal system of Vietnamese. Such votes are not necessarily "incorrect," though they may be morally dubious. But let us imagine a worse case than any of those.

The city of Cambridge, Massachusetts, as of the fall of 1983, did not yet have cable television, nor had it made any contract for the installation of cable TV. I have somewhere read that there is disagreement as to whether Cambridge should itself undertake the installation or whether it should contract with an outside company. That is as much as I know. Suppose that by some Satanic maladvertence I were appointed to the group charged with making the decision. Suppose further that all members of the group were unknown to me, that I could not evaluate their professional credentials, that when the discussion took place the arguments made were all completely beyond my comprehension, and that the group, quite unaware of my total incompetence, expected me to cast a vote for either city control of cable or outside control. I ought, of course, to excuse myself and go home, but maybe I do not want to admit to my full ignorance. Would there be any way that I could, listening to a discussion of which I understood nothing, guess at the correct answer? In fact, there would be.

In a good expressive discussion, such as the deliberations in *Twelve Angry Men*, information is transmitted that is not part of any relevant argument. Each participant in some degree expresses his individuality: assertiveness, intelligence, education, fair-mindedness, compassion, prejudice, cowardice, and so on. Each participant also reveals which side of the issue he supports and how strongly. In such circumstances, positions on the issue *take on meanings* even if the issue itself is not understood at all. In the imaginary case of Cambridge cable, I might learn that the position labeled "city control" is championed by people who seem to be authoritarian, not highly educated, distrustful of others, and identified with wealth, whereas the position "outside control" is championed by people who seem to me to be fair-minded, highly educated, and equally concerned for the welfare of all citizens. The casting of my ignorant vote becomes a pure act of self-presentation. What sort of person do I myself want to be, and what sort of person do I want to appear to be? Not knowing the cognitively correct answer, there is still an answer correct for me; it is the answer championed by the kind of people I admire and wish to resemble. (Remember, please, that the suggested linkage between position and personality is entirely imaginary. I really do not know anything about the issue.)

In an important way the extreme case imagined is like social comparison as it operates in risk problems: The individual has certain values he wishes to real-

ize, to express in his decision, but he cannot know how to realize his values until he finds out the distribution of positions in the group. In another way the case imagined is unlike decisions on risk or attitudes on race or gender: The individual knows nothing about the positions and so cannot make an initial prediscussion decision except by tossing a coin. The choice dilemmas offering various odds options or seven-point-scale positions, and also the scale on prejudice, contain some information, enough to give the individual the strong impression that he knows in advance of discussion how to make a decision congruent with his values. If he wants to be a risk-taker or unprejudiced, he at least thinks he knows on which side of the middle of the scale he ought to place himself. The group polarization phenomenon, setting aside persuasive arguments, indicates that people do not know everything they need to know in order to express the values they hold until a real distribution of positions becomes public.

The jurymen in Rose's play have as their decision options "guilty" and "not guilty." In advance of the trial and of the jury deliberations, those options have no meaning at all, either cognitive or expressive. Then comes a trial that the judge says has been long and complex. Now, assembled in the juryroom, each individual has information on which to base a decision, and each has made a decision. However, there is disagreement within the group. Originally, only one member dissents, but after some discussion, he is joined by another and later another, and by the time they take a second formal ballot, opinions are evenly divided: Six vote not guilty and six vote guilty. How is an individual juryman to know what to do? He ought, of course, to attend to the relevant arguments coming from either side and let his decision be determined by the weight of argument. However, the arguments are many and complex, and so it is possible that some attention would be paid to social comparison as an alternative route to the right decision, a route less demanding than the purely intellectual one.

How would a conscientious juror wish to present himself to himself and to others? Presumably as a person able intellectually to evaluate the evidence, as disinterested and conscientious, as strong enough to convict if necessary but compassionate enough to acquit if there is reasonable doubt. How can he act so as to seem to have all those virtues? The hard way is, in fact, to have them all, but not everyone can manage that. There is an easier way, and that is to vote as those vote who seem to be intellectually capable, disinterested, conscientious, strong, and compassionate. How can one find out who these people are? By the way they conduct themselves in the discussion, by all their expressive behavior, by the great residue of action that goes beyond statement of relevant arguments.

The published form of Rose's play has an odd feature. Each juryman, named only by number, is characterized by the author, before the play starts, in a single-paragraph vignette. Here are the full vignettes for Juror No. 8 (Henry Fonda) and his opposite number, the juror most strongly attached to conviction, Juror No. 3 (Lee J. Cobb in the movie):

Juror No. Eight: He is a quiet, thoughtful, gentle man—a man who sees all sides of every question and constantly seeks the truth. He is a man of strength tempered with compassion. Above all, he is a man who wants justice to be done, and will fight to see that it is. [p. 5]

Juror No. Three: He is a very strong, very forceful, extremely opinionated man within whom can be detected a streak of sadism. Also he is a humorless man who is intolerant of opinions other than his own, and accustomed to forcing his wishes and views upon others. [p. 4]

When you read the two character vignettes and know that No. 8 is the juror most attached to "not guilty" and No. 3 the juror most attached to "guilty," the response options take on a meaning. It is a meaning independent of any relevant arguments, for none has been heard when one reads the vignettes. It is a meaning that results from the implicit question: "Who would I prefer to agree with?"

As it happens, the author has built into his play, not deliberately I would guess, but intuitively, an almost perfect correlation between the social desirability of each juror and the degree to which each juror is attached to conviction or acquittal. Here are the full vignettes for, respectively, the juror next to No. 8 in his attachment to acquittal (No. 9, the old man) and the juror next to No. 3 in his attachment to conviction (No. 10):

Juror No. Nine: He is a mild, gentle old man, long since defeated by life, and now merely waiting to die. He recognizes himself for what he is, and mourns the days when it would have been possible to be courageous without shielding himself behind his many years. [p. 5]

Juror No. Ten: He is an angry, bitter man—a man who antagonizes almost at sight. He is also a bigot who places no value on any human life save his own. Here is a man who has been nowhere and is going nowhere and knows it deep within him. [p. 5]

It becomes increasingly clear that "not guilty" is the good guys' ticket, as it were, and "guilty," the bad guys'. The little vignettes written as guides to the actors suggest no great characterological complexity, but then this is not a play of character; it is a play almost purely about group dynamics.

How can one determine from the action of the play the degree of adherence of each character to the alternative verdicts? There are two ways. The main way is the order in which jurymen switch from their initial vote for "guilty" to an eventual vote for "not guilty." Some of these changes of heart occur in one person at a time and some occur in two or three at once—as on the second formal ballot. Because they do not all move one at a time, there are some ties in adherence to a verdict, and so only a partial ordering is possible. Using hyphens to link jurors who switch at the same time and semicolons to separate jurors who switch at different points, the partial ordering by vote change from "not guilty" (anchored by Henry Fonda, No. 8, who votes that way from the start) is:

8; 9; 5; 2-6-11; 7-12-Foreman; 10-4; 3.

Adherence to one verdict or the other is also revealed, rather astonishingly, by the order in which the author directs that hands be raised on the initial ballot, voting "guilty." Using hyphens to connect simultaneous votes and semicolons to separate votes that can be ordered, the order of adherence from "guilty" now to "not guilty" (the reverse of the vote switches) is: 3-10-7-12; 4-Foreman-2-6-5; 11; 9; 8. It is, again, the case that No. 8 is distinctive because he never does vote "guilty"; he, therefore, anchors, as on the vote switches, the "not-guilty" extreme. Arraying the two orders from "not guilty" to "guilty," it becomes clear that we have two indices of the same thing:

Verdict switches: N.G. 8; 9; 5; 2-6-11; 7-12-Foreman; 10-4; 3
Voting order: N.G. 8; 9; 11; 2-6-5-Foreman-4; 3-10-12-7

In Figure 6-5 the jurors (identified by number) are arrayed from left to right in the more differentiated order, the order of verdict switches. I take it for granted that you will agree that social desirability declines from left to right.

The vignettes are advance guidelines for actors, but the jurors do not hold up cards describing their characteristics. How are the characteristics revealed? In dialogue and action strewn across three acts but generally extremely unambiguous fairly early on. Here are some characterizing speeches:

JUROR NO. 8: I want to talk for a while. Look—this kid's been kicked around all his life. You know—living in a slum—his mother dead since he was nine. That's not a very good head start . . . [p. 5]
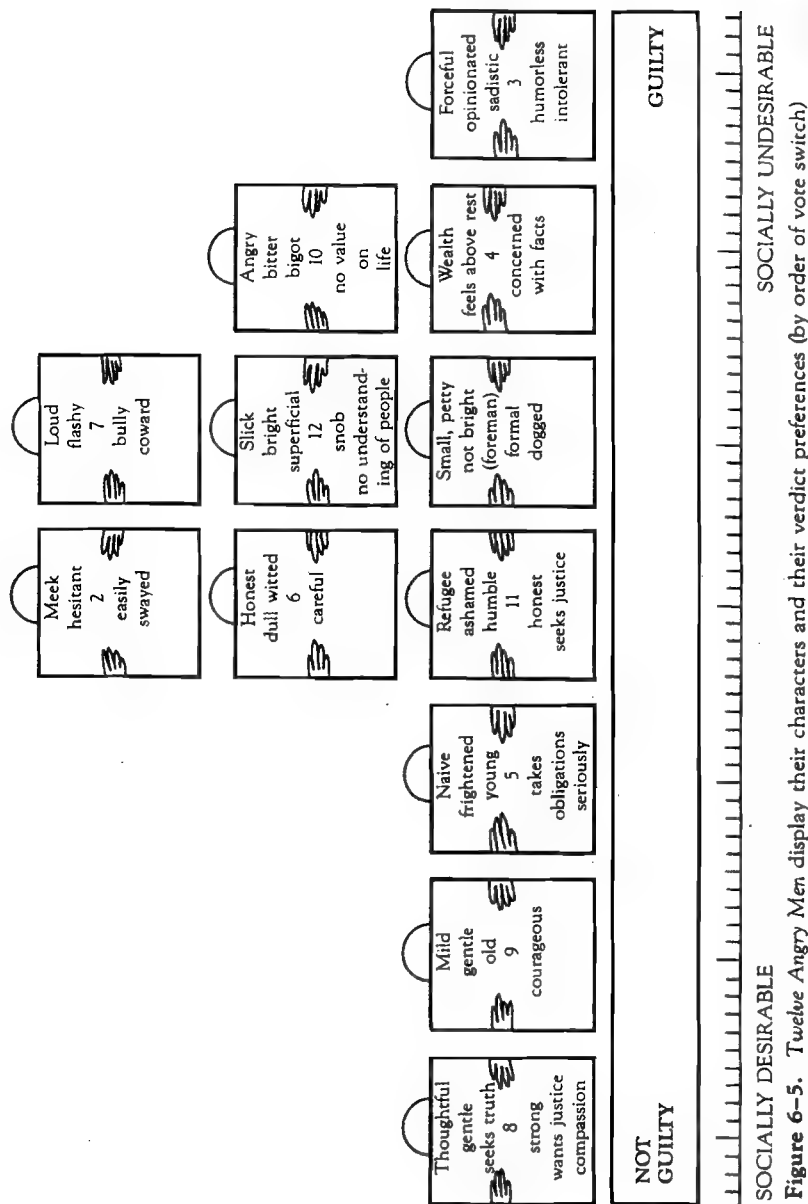
Register "compassionate." Here is an early speech of his opposite number:

JUROR NO. 3: You're right. It's the kids. The way they are—you know? They don't listen (Bitterly). I've got a kid. When he was eight years old he ran away from a fight. I saw him. I was so ashamed. I told him right out, "I'm gonna make a man out of you or I'm gonna bust you up into little pieces trying." When he was fifteen he hit me in the face. He's big, you know? I haven't seen him in three years. Rotten kid! I hate tough kids! [p. 5]

Not exactly disinterested and certainly disposed to force his ways on others.

The jurors do not hold up cards on which their characters are written, but imagine a Theater of the Absurd in which they would (Figure 6-5). It would certainly make social comparison easy. And, given that one must choose between just two options, "guilty" and "not guilty," can there be any doubt which is the "right" answer? Once the correlation between virtue and verdict emerges, any doubtful juror, having trouble weighing the relevant arguments, would know how to vote.

The near-perfect correlation between attraction to acquittal and social desirability is the second unlikely premise the playwright asks us to entertain. Juror No. 4 is the only person seriously out of line, and he has to be—for a reason of dramaturgy. The vignette reads as follows.

SOCIALLY DESIRABLE

GUILTY

Forceful opinionated sadistic 3 humorless intolerant

Angry bitter bigot 10 no value on life

Wealth feels above rest 4 concerned with facts

Loud flashy 7 bully coward

Slick bright superficial 12 snob no understanding of people

Small, petty not bright (foreman) 1 formal dogged

Meek hesitant 2 easily swayed

Honest dull witted 6 careful

Refugee ashamed humble 11 honest seeks justice

Naive frightened young 5 takes obligations seriously

Mild gentle old 9 courageous

Thoughtful gentle seeks truth 8 strong wants justice compassion

NOT GUILTY

SOCIALLY DESIRABLE

SOCIALLY UNDESIRABLE

Figure 6-5. *Twelve Angry Men* display their characters and their verdict preferences (by order of vote switch)

Juror No. Four: He seems to be a man of wealth and position and a practiced speaker who presents himself well at all times. He seems to feel a little bit above the rest of the jurors. His only concern is with the facts in this case and he is appalled with the behavior of the others. [p. 4]

Juror No. 4 is especially appalled by his closest neighbors in Figure 6-5: a sadistic swine (12) and a bitter bigot (10). Those two are not equipped to keep the play going by making persuasive arguments in a temperate tone, designed to counter the arguments of No. 8. Gentlemanly, reasonable No. 4, however, is able to sustain some uncertainty, and it is only when he is persuaded—on the evidence of the eyeglasses—that the far right collapses.

The first premise the playwright implicitly asks us to entertain is the possibility that a person accused of murder could be so inadequately defended at his trial as to yield a near unanimous verdict that is incorrect as a representation of the total pool of arguments. Granted, the premise the "improbable" vote switch in the course of the play follows quite strictly from the theory of persuasive arguments. The second premise is the high correlation between attraction to acquittal and social desirability. Granted that premise, the vote switch follows quite strictly from the theory of social comparison. The author has proved *his* case in two ways, just the two that the theory of group polarization provides. He cleverly leaves implicit and barely detectable the two improbabilities he does not attempt to make credible—the unrepresentative case and the virtue–verdict correlation. The improbability that cannot be missed, the improbability that gives the play its punch, is the vote switch, and that improbability is rendered fully credible. While the play's construction is not exactly a proof in geometry, it has some of the elegance of such a proof.

Scientific Jury Selection

If group polarization operates in the jury room, it would seem to follow that selection of the jury is an all-important process. That must be so, because except in the improbable case of an unrepresentative trial, like the one in *Twelve Angry Men*, the initial inclination of the jury simply becomes more extreme in group discussion. The trick, then, from an attorney's point of view, whether prosecuting or defending, is to get jurors with the "right" initial inclination. There is nothing new in this, and books on trial tactics are filled with guides to juror selection. What is new and is sometimes regarded as a strong threat to our judicial system and, less often, as a major advance in attaining justice, is the use since 1972 of a technology that promises greatly to improve upon intuitive guides for selecting juries—even those of such eminent attorneys as F. Lee Bailey and Melvin Belli.

The idea is simply to do an opinion survey of the community from which jurors will be drawn, asking respondents how they would be inclined to vote, for example, in the case of a particular defendant X accused of murdering a par-

ticular Y. Those pretrial inclinations are then related to age, sex, political affili-
ation, marital status, and every sort of personality characteristic. From the cor-
relations between opinions on the trial and juror characteristics, it is possible to
work out profiles of the "good" juror and the "bad" juror from the point of
view of the side that commissioned the survey. Then in the *voir dire* (Old
French: to speak the truth) procedure, in which prospective jurors are interro-
gated, an attorney can use what are called his peremptory challenges (free dis-
missals made without having to show cause) to eliminate the "bad" jurors and
so incline the jury his way.

The process of jury selection begins with some kind of list of membership in
a community, most often voter registration rolls, but because young people,
minorities, and the poor are underrepresented on those rolls, they are some-
times supplemented with welfare lists or lists of persons holding driver's licen-
ses. The Supreme Court has, in a series of decisions, taken the position that a
jury must be a representative cross-section of the community in which the
offense was committed. The reasons for doing so are, from a group problem-
solving point of view, excellent. "Representative" has, in practice, generally
meant heterogeneous, and problems are in general more likely to be solved if a
variety of perspectives and experiences are brought to bear upon them.

From the source list of members of a community the jury commissioner's
staff selects (usually at random) a pool of potential jurors, who are sent a ques-
tionnaire. Some are disqualified (not knowing English is a clear disqualifica-
tion), and some are exempted (doctors, nurses, mothers of infants, and so on),
but eventually a pool of prospective jurors, called the venire, is selected to serve
for a given period, usually about a month. Then comes the *voir dire* procedure
in which defense and prosecuting attorney and/or the judge question members
of the venire to ascertain whether or not they can serve without bias. A panel
member, a potential juror, may be excused in either of two ways. A "challenge
for cause" occurs when the judge is persuaded that a panel member cannot be
impartial. In addition, both defense and prosecution are allowed some number
(e.g. six to twelve) of "peremptory challenges," for which it is not necessary to
give any reason. The number of peremptory challenges varies with the jurisdic-
tion, the seriousness of the charge, and the number of defendants.

In trials with political implications, such as the trials on conspiracy charges
in the late 1960s and 1970s of Angela Davis, Black Panther Huey Newton, the
Berrigan brothers and, in the post-Watergate period, of John Mitchell and
Maurice Stans, the probability of bias was very great, and the *voir dire* was
lengthy and searching. The number of peremptory challenges allowed the
defense was also unusually large. Lawyers have for centuries tried to select jur-
ors favorable to their side of a case, and they quite naturally and inevitably
have developed generalizations about how to do so. It is equally natural and
inevitable that those generalizations have never been impartially tested and
that most of them simply traffic in stereotypes. Bailey and Rothblatt, in their
Successful Techniques for Criminal Trials (1971), advise defense lawyers to choose

women if the principal witness against the defendant is female, because women
are "somewhat distrustful" of other women. On the question of wealth and sta-
tus, Clarence Darrow (1936) warns that the wealthy will convict unless the
defendant is accused of a white-collar crime. Goldstein (1935) has ranked eth-
nic groups on emotionalism, from high to low: Irish, Jewish, Italian, French,
Spanish, Slavic, and Nordic. The last (including English, Scandinavians, and
Germans) are to be preferred if it will be necessary to combat emotional
appeals.

The examples just given probably do not inspire much confidence in the
juror selection tactics that lawyers have developed by unsystematic trial and
error, but of course there are also some more prepossessing examples. It would
be extremely interesting to know how much truth there is in each general-
ization that has been put forward, but no one is ever going to carry out that
research program.

In the winter of 1971–72, the Harrisburg Seven, including the Berrigan
brothers, were tried for conspiracy to raid draft boards and destroy records;
blow up the heating tunnels in Washington, D.C.; and kidnap Secretary of
State Henry Kissinger. The trial was held in Harrisburg, Pennsylvania, a very
conservative and pro-government community, and so a group of sympathetic
social scientists joined the defense team, headed by former Attorney General
Ramsey Clark. That group made the first methodologically sound use of the
technique called "scientific jury selection."

They carried out surveys of samples in Harrisburg, asking about attitudes
toward the defendants and their case, administering various sorts of possibly
relevant personality scales and collecting such demographic data as age, sex,
occupation, race, education, media contacts, and hobbies. There is really no
telling what might be significantly related to attitudes toward a case. One might
find young, egalitarian females who read *The New York Times* well disposed to
the defense and older, authoritarian males who read the *New York Daily News*
on the side of the prosecution. Those data were combined (nowadays it would
be by computer) to yield profiles of the good juror and the bad juror. Then, in
the *voir dire*, defense attorneys asked such seemingly "innocent" questions as,
"What newspapers do you read regularly?" and used their peremptory challen-
ges (several hundred were granted) to excuse bad jurors. The trial ended in a
hung jury, 10–2 favoring the defense. A clear victory for the defense. Is it also
clear proof that scientific jury selection works? Not quite.

For one thing, the Harrisburg defense did not rely entirely on its surveys
and profiles; in fact, the social scientists, not knowing whether they would
work well or not, advised against such total reliance, so the lawyers also exer-
cised their own judgment. The story is that the computer had vetoed one juror
that Ramsey Clark decided to keep, and that juror was one of the two who held
out for conviction—that too, however, is not serious evidence. In addition to
the informed intuition of the lawyers, the defense made use of an old technique
used to some extent also by prosecution. While it is illegal directly to approach

a venireman, it is not illegal to ask around about him; in short, you can ask someone who knows someone on the venire. That is a rather tricky thing to do, because if the target gets wind of your inquiries, he may develop a strong bias against you (Berman and Sales, 1977; Bonora and Krauss, 1979; Kairys, Schulman, and Harring, 1975; McConahay, Mullin, and Frederick, 1977).

Scientific jury selection has created a great stir; there are commercial firms that will do the whole job for a rather large fee. And here lies a possible danger to our justice system since only wealthy defendants (usually corporations) may be able to afford scientific jury selection. The curious thing is that the faith in the method is not supported by the best research on the topic (Berman and Sales, 1977; Hastie, Penrod, and Pennington, 1983; Penrod, 1979; Saks, 1977). What convinces many is the very high batting average. Almost no cases have been lost by a defense that used the method. However, until recently (the John DeLorean drug case and the Agent Orange class action suit) these cases have usually been highly political cases where the charge was conspiracy, and conspiracy is a difficult charge to prove. As it turns out (Hans and Vidmar, 1982; Saks and Hastie, 1978), the records of cases that did not use scientific jury selection but are comparable to those that did are equally good. So the high batting average is just no evidence at all—which does not necessarily mean that defendants with plenty of money will neglect to use it.

The work that has been done to assess the value of scientific jury selection is actually heavily loaded in favor of the method. The practice has been to determine many individual characteristics of actual jurors (always in simulated juries) and to find out how well those characteristics predict verdict preferences on a first, posttrial predeliberation verdict. This problem is, in two ways, easier than the problem that confronts any actual practitioner. In the first place, he is not allowed to approach actual jurors but must generalize to them from the study of samples of populations to which jurors belong. A more serious difficulty is that no survey can present respondents with the real problem jurors will face, which is forming verdict preferences *after* hearing a trial. A trial includes evidence, testimony, and argument that may last some days, and everything indicates (Hans and Vidmar, 1982; Hastie, Penrod, and Pennington, 1983; Saks, 1976) that the content of the trial is much the most important determinant of initial verdict preferences. The respondent to a survey cannot be put in the juror's posttrial position. All the survey can do is try to find a question highly relevant to or predictive of verdict preferences, and such questions are hard to find.

The problem that has been posed in research designed to evaluate scientific jury selection is much easier than the field problem, and so if it were solved, we should not really know whether the field problem can be solved. However, if even the much-simplified procedure of predicting posttrial verdict preferences from individual characteristics of jurors does not work, then we can be sure that the more difficult field procedure will not work. A single example will serve

represent a strikingly unanimous research outcome (Berman and Sales, 1977; Hastie, Penrod, and Pennington, 1983; Penrod, 1979; Saks, 1977).

In the highly realistic simulation study of Hastie, Penrod, and Pennington (1983), the 828 participating jurors provided information, including age, gender, occupation, residence, education, political party, marital status, race, income, and number of previous cases heard as a juror. That information was entered into a multiple regression equation designed to combine predictors in the most powerful way with posttrial, predeliberation verdict preferences as the dependent variable. The result: Only four predictors had any significant relation to verdict preferences, and the total R was .179, which is no use at all. A subsample of 269 jurors completed a more extensive questionnaire that included information about reading habits, attitude toward the death penalty, and other characteristics. Just five items had any relation to verdict preferences. Those were optimally combined, and verdict preferences simply dichotomized into votes favoring conviction for murder and votes favoring acquittal. In this form the votes of 61 percent of the jurors were correctly predicted. The result is a bit better than chance, but it hardly suggests a powerful technique. Furthermore, it must be a considerable overestimate of what can be accomplished in a real situation.

Even though the weight of the experimental evidence is that scientific jury selection is not of much use (at least in felony trials), I would not expect the market for the service to disappear, mainly because the idea seems too good. How can it fail eventually to work to some degree? In 1984 some "decision research" companies, as they call themselves, do not simply sample attitudes in advance of the real trial. They conduct advance mock trials, many times, before juries that are representative of the juror pools, and instead of simply trying to predict final votes from demographic characteristics they interview jurors on the main aspects of the trial.

If scientific jury selection ever attains substantial demonstrated power, there are many ways in which that power could be and probably would be checked. There is, in the first place, considerable risk in carrying out a prolonged *voir dire* with many questions not clearly relevant to the detection of bias; the attorney may prejudice the jury against his case if he seems to be trying to do what he may be trying to do, select a jury partial to his client. There are some attorneys even today who think the best strategy to use in the *voir dire* is unconcerned acceptance of the first twelve possible jurors, because the expressive meaning of this stance is roughly that of having so clear a case that it can safely be entrusted to anyone at all, and that expressive meaning is more important than the detection of bias (Blunk and Sales, 1977). In the second place, all questions in the *voir dire* are supposed to be designed to detect bias, and the judge can rule out any questions that do not seem to him to have this purpose. Finally, of course, the judge can, and in Massachusetts often does, simply relieve the attorneys of the interrogation task and take it entirely

upon himself. What then would be the use of profiles of good jurors and bad jurors?

Summary

In the play *Twelve Angry Men* an initial juror ballot of 11–1 in favor of "guilty" becomes changed by jury discussion to a final verdict unanimously favoring "not guilty." Such an outcome would be an extreme statistical rarity, and yet in the play it seems entirely credible. That is because the play makes the outcome credible by means of the expert deployment of two processes of group dynamics: persuasive arguments and social comparison. The audience is led to believe that an extremely unrepresentative trial has taken place, with a completely inadequate defense, and that makes it possible for persuasive arguments to lead to a correct verdict in what amounts to a new and more representative trial held in the jury room. The audience accepts, without noticing it, an almost perfect correlation between the character and competence of twelve jurymen and their degree of attraction to a "not guilty" verdict. This makes it possible for social comparison processes to explain the "unrealistic" vote switch of the play.

Conceptualization of the two basic group processes—persuasive arguments and social comparison—was developed over a period of two decades, beginning in 1962. Development was consistently in the direction of increased generality and abstraction. The processes were first thought of as explanations of the risky shift, then as explanations of shifts either to increased risk *or* increased caution, then as processes explanatory of group polarization with respect to any content whatever—including trials at law. Group polarization is the following phenomenon: The means of individual decisions or attitudes, if they fall to one side of psychological neutrality, will, following group discussion by like-minded persons, become more extreme or polarized in the direction of the original inclination.

The theory of persuasive arguments (one of two sufficient explanations of group polarization) has the following essential features:

1. For each decision or judgment or attitude position, there exists a (culturally and historically relative) pool of arguments pro and con, with the balance varying from item to item.
2. Each person who individually takes a position draws upon some sample of the total pool of arguments; that sample determines the position he takes; and the character of the pool determines the mean of the individual positions. Prior to discussion, the arguments considered by any set of individuals are only partially shared.
3. In discussion, arguments are expressed and shared, and so each individual takes account of a larger and more representative portion of the total pool, and decisions shift toward the position anchored by the total

pool, which will be away from psychological neutrality in the same direction as the mean of the initial decisions, but more extreme or polarized.

The theory of social comparison (a second sufficient explanation of group polarization) has the following essential features:

1. Any given problem, decision, or dilemma engages in most people the desire to express certain values or virtues.
2. In advance of information as to the distribution of positions taken by other persons, each imagines that his own position does, in fact, express the relevant values or virtues.
3. When, in discussion, the actual distribution of positions is made known, those whose positions do not express the relevant values will be motivated to shift.

There is reason to think that group polarization operates in the deliberations of real juries; 90 percent of juries that do not hang reach verdicts consistent with the original ballot and, by definition, more extreme than that ballot. Because jury deliberation primarily amplifies initial jury inclination, it would seem to be very important for the winning of a case to have the right jury. A method of scientific jury selection has been invented which, in essence, tries to predict juror inclinations on a case from individual characteristics shown by empirical survey methods to be related in the general community. There are two deep difficulties. The problem a real jury considers is nothing less than the total hearing of a case plus their own group deliberations upon it. That problem cannot be presented by survey methods to a community sample, and the surrogate problem (a questionnaire) may not realistically represent what the people surveyed would do if required to serve. Furthermore, it seems not to be the case that jurors' predeliberation verdict preferences are determined by demographic characteristics or prior attitudes. It looks very much as if jurors try to carry out their solemn duty as impartial fact finders and so establish verdict preferences based mainly on the trial evidence.

References

ANDERSON, N. H., and C. C. GRAESSER. 1976. An information integration analysis of attitude change in group discussion, *Journal of Personality and Social Psychology*, 34: 210–22.

BAILEY, F. L., and H. B. ROTHBLATT. 1981. *Successful Techniques for Criminal Trials*. New York: Lawyers Cooperative.

BERMAN, J., and J. B. SALES. 1977. A critical evaluation of the systematic approach to jury selection, *Criminal Justice and Behavior*, 4: 219–40.

BISHOP, G. D., and D. G. MYERS. 1974. Informational influence in group discussion, *Organizational Behavior and Human Performance*, 12:902–1104.

BLUNK, R. A., and B. D. SALES. 1977. Persuasion during the *voir dire*. In B. D. Sales (ed.), *Psychology in the Legal Process*. New York: Spectrum, pp. 39–58.

BONORA, B., and E. KRAUSS. 1979. *Jury Work: Systematic Techniques*. National Jury Project.

BROWN, R. 1965. *Social Psychology*. New York: Free Press.

BURNSTEIN, E. 1982. Persuasion as argument processing. In H. Brändstetter, J. H. Davis, and G. Stocker-Kreichgauer (eds.), *Group Decision Making*. London: Academic Press, pp. 103–24.

BURNSTEIN, E., and A. VINOKUR. 1973. Testing two classes of theories about group induced shifts in individual choice, *Journal of Experimental Social Psychology*, 9:123–37.

——. 1975. What a person thinks upon learning he has chosen differently from others: Nice evidence for the persuasive arguments explanation of choice shifts, *Journal of Experimental Social Psychology*, 11:412–26.

CLARK, R.; W. H. CROCKETT; and R. L. ARCHER. 1971. Risk as value hypothesis: The relationship between perception of self, others, and the risky shift, *Journal of Personality and Social Psychology*, 20: 425–29.

CODOL, J.-P. 1975. On the so called "superior conformity of the self" behavior: Twenty experimental investigations, *European Journal of Social Psychology*, 5: 457–501.

DARROW, C. 1936. Attorney for the defense, *Esquire Magazine*, May.

DAVIS, J. H.; N. L. KERR; R. S. ATKIN; R. HOLT; and D. MEEK. 1975. The decision processes of 6- and 12-person mock juries assigned unanimous and two-thirds majority rules, *Journal of Personality and Social Psychology*, 32: 1–14.

DOISE, W. 1969. Intergroup relations and polarization of individual and collective judgments, *Journal of Personality and Social Psychology*, 12: 136–43.

EBBESEN, E. B., and R. J. BOWERS. 1974. Proportion of risky to conservative arguments in a group discussion and choice shift, *Journal of Personality and Social Psychology*, 29: 316–27.

FESTINGER, L. 1954. A theory of social comparison processes, *Human Relations*, 7: 117–40.

FRASER, C. 1971. Group risk-taking and group polarization, *European Journal of Social Psychology*, 1: 493–510.

FRASER, C.; C. GOUGE; and M. BILLIG. 1971. Risky shifts, cautious shifts, and group polarization, *European Journal of Social Psychology*, 1: 7–30.

GOLDSTEIN, I. 1935. *Trial Technique*. Chicago: Callaghan.

HANS, V. P., and N. VIDMAR. 1982. Jury selection. In N. L. Kerr and R. M. Bray (eds.), *The Psychology of the Courtroom*. New York: Academic Press, pp. 39–82.

HASTIE, R.; S. PENROD; and N. PENNINGTON. 1983. *Inside the Jury*. Cambridge, Mass.: Harvard University Press.

HINDS, W. C. 1962. Individual and group decisions in gambling situations. Unpublished master's thesis, School of Industrial Management, Massachusetts Institute of Technology.

JELLISON, J. M., and J. RISKIND. 1970. A social comparison of abilities interpretation of risk-taking behavior, *Journal of Personality and Social Psychology*, 15: 375–90.

——. 1971. Attribution of risk to others as a function of their ability, *Journal of Personality and Social Psychology*, 20: 413–15.

ELLISON, J. M.; J. RISKIND; and L. BROLL. 1972. Attribution of ability to others on skill and chance tasks as a function of level of risk, *Journal of Personality and Social Psychology*, 22: 135–38.

KAIRYS, D.; J. SCHULMAN; and S. HARRING (eds.). 1975. *The Jury System: New Methods for Reducing Prejudice*. Philadelphia: National Jury Project and National Lawyers Guild.

KALVEN, H., JR., and H. ZEISEL. 1966. *The American Jury*. Boston: Little, Brown.

KAPLAN, M., and G. D. KEMMERICK. 1974. Juror judgment as information integration: Combining evidential and nonevidential information, *Journal of Personality and Social Psychology*, 30: 493–99.

KERR, N. L., and R. J. MACCOUN. 1985. The effects of jury size and polling method on the process and product of jury deliberation, *Journal of Personality and Social Psychology*, 48: 349–63.

LAMM, H., and D. G. MYERS. 1978. Group-induced polarization of attitudes and behavior. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*. New York: Academic Press, 11: 145–95.

LEVINGER, G., and D. J. SCHNEIDER. 1969. Test of the "risk is a value" hypothesis, *Journal of Personality and Social Psychology*, 11: 165–69.

McCONAHAY, J.; C. MULLIN; and J. FREDERICK. 1977. The uses of social science in trials with political and racial overtones: The trial of Joan Little, *Law and Contemporary Problems*, 41: 205–29.

MOSCOVICI, S., and M. ZAVALLONI. 1969. The group as a polarizer of attitudes, *Journal of Personality and Social Psychology*, 12: 125–35.

MYERS, D. G. 1978. Polarizing effects of social comparison, *Journal of Experimental Social Psychology*, 14: 554–63.

——. 1982. Polarizing effects of social interaction. In M. Brandstätter, J. H. Davis, and G. Stocker-Kreichgauer (eds.), *Group Decision Making*. London: Academic Press, pp. 125–61.

MYERS, D. G., and S. J. ARENSON. 1972. Enhancement of dominant risk tendencies in group discussion, *Psychological Reports*, 30: 615–23.

MYERS, D. G.; P. J. BACH; and F. B. SCHREIBER. 1974. Normative and informational effects of group interaction, *Sociometry*, 37: 275–86.

MYERS, D. G., and G. D. BISHOP. 1971. The enhancement of dominant attitudes in group discussion, *Journal of Personality and Social Psychology*, 20: 386–91.

MYERS, D. G., and M. F. KAPLAN. 1976. Group-induced polarization in simulated juries, *Personality and Social Psychology Bulletin*, 2: 63–66.

MYERS, D. G., and H. LAMM. 1976. The group polarization phenomenon, *Psychological Bulletin*, 83: 602–27.

NORDHØY, F. 1962. Group interaction in decision-making under risk. Unpublished master's thesis, School of Industrial Management, Massachusetts Institute of Technology.

PADAWER-SINGER, A.; A. N. SINGER; and R. L. J. SINGER. 1977. An experimental study of twelve vs. six member juries under unanimous vs. nonunanimous decisions. In B. D. Sales (ed.), *Psychology in the Legal Process*. New York: Spectrum, pp. 77–81.

PENROD, S. D. 1979. Study of attorney and "scientific" jury selection models. Unpublished doctoral dissertation, Harvard University.

PRUITT, D. G. 1971a. Choice shifts in group discussion: An introductory review, Journal of Personality and Social Psychology, 20: 339-60.

———. 1971b. Conclusions: Toward an understanding of choice shifts in group discussion, Journal of Personality and Social Psychology, 20: 494-510.

ROSE, R. 1955. Twelve Angry Men. Chicago: Dramatic Publishing Co.

SAKS, M. J. 1976. The limits of scientific jury selection: Ethical and empirical, Jurimetrics Journal, 17: 3-22.

———. 1977. Jury Verdicts. Lexington, Mass.: Heath.

SAKS, M. J., and R. HASTIE. 1978. Social Psychology in Court. New York: Van Nostrand Reinhold.

SANDERS, G. S., and R. S. BARON. 1977. Is social comparison irrelevant for producing choice shifts? Journal of Experimental Social Psychology, 13: 303-14.

STOKES, J. P. 1971. Effects of familiarization and knowledge of others' odds choices on shifts to risk and caution, Journal of Personality and Social Psychology, 20: 407-12.

STONER, J. A. F. 1961. A comparison of individual and group decisions including risk. Unpublished master's thesis, School of Management, Massachusetts Institute of Technology.

———. 1968. Risky and cautious shifts in group decisions: The influence of widely held values, Journal of Experimental Social Psychology, 4: 442-59.

TANFORD, S., and S. PENROD. 1983. Computer modeling of influence in the jury: The role of the consistent juror, Social Psychology Quarterly, 46: 200-12.

TEGER, A. I., and D. G. PRUITT. 1967. Components of group risk taking, Journal of Experimental Social Psychology, 3: 189-205.

VALENTI, A. C., and L. L. DOWNING. 1975. Differential effects of jury size on verdicts following deliberation as a function of the apparent guilt of a defendant, Journal of Personality and Social Psychology, 32: 655-63.

VINOKUR, A. 1971. Cognitive and affective processes in influencing risk taking in groups: An expected utility approach, Journal of Personality and Social Psychology, 20: 472-86.

VINOKUR, A., and E. BURNSTEIN. 1974. Effects of partially shared persuasive arguments on group-induced shifts: A group-problem-solving approach, Journal of Personality and Social Psychology, 29: 305-15.

WALLACH, M. A., and N. KOGAN. 1959. Sex differences and judgment processes, Journal of Personality, 27: 555-64.

———. 1961. Aspects of judgment and decision making: Interrelationships and changes with age, Behavioral Science, 6: 23-36.

WALLACH, M. A.; N. KOGAN; and D. J. BEM. 1962. Group influence on individual risk taking, Journal of Abnormal and Social Psychology, 65: 75-86.

WALLACH, M., and C. W. WING, JR. 1968. Is risk a value? Journal of Personality and Social Psychology, 9: 101-6.

WHYTE, W. H., JR. 1956. The Organization Man. New York: Simon & Schuster.

III

Some Psycholegal Issues

AMERICAN LAW TODAY enjoys the invigorating impact of a social psychology that questions its every assumption and promises eventually to motivate reforms. Social psychology, in turn, benefits from a sharp cross-examination, from the legal perspective, of its methods and findings and gains the gratification of feeling relevant to the good of society. The law at first made the mistake of citing the conclusions of studies without evaluation of research design, very much as if psychological studies were legal precedents. Social psychologists at first made the mistake of thinking that lawyers, judges, and jurors would be as ready as we are to believe that results obtained from college undergraduates will hold true for humans everywhere. The law has been sensitized to research design or internal validity: Is it clear that variation in dependent variables is caused by independent variables? Social psychology has been sensitized to external validity: Is it clear that findings obtained in one setting—the laboratory—will hold true in another—the arraignment to set bail or the jury room (Carlsmith, Ellsworth, and Aronson, 1976)?

It took a while for the law to take notice of social psychology, but it cannot disregard the field today, because every year brings some result that threatens to dynamite a foundation. In 1979, for instance, Kassin and Wrightsman found that while the judge's instructions to the jury are typically given at the end of the trial, when the evidence has all been heard, they might better be given at the start, before the evidence is heard. In 1980 Grisso and Manoogian found that the majority of "juveniles" are not able to understand the Miranda warn-

ing advising them of their rights in a police interrogation. In 1981 Thompson, Fong, and Rosenhan found that jurors do not do what the judge directs them to do with respect to inadmissible evidence, which is to strike it from their memories as it is stricken from the written record. In 1982 Wissler and Saks found that although jurors are supposed to disregard a defendant's prior criminal record in assessing the likelihood of guilt, they are very far from doing so. In 1983–84 Phoebe Ellsworth and her associates (1984) put together massive evidence that jurors willing to impose a death penalty and so "death-qualified" in capital cases are more prone to agree with the prosecution and convict than are potential jurors opposed to the death penalty and so "excludable" from juries in capital cases (Cowan, Thompson, and Ellsworth, 1984; Fitzgerald and Ellsworth, 1984). This means that those willing to impose the ultimate penalty in addition tend to be biased against the defendant.

We are social psychologists, not law students, and so the problems that researchers have chosen to study most deeply, and the problems that I have chosen to examine here, have been selected not because they are the most frequent events in the justice system nor even because they are the most consequential. Eyewitness identification (Chapter 7) and jury selection (Chapter 8) have been given special attention because they are applied instances of classic topics in social psychology, instances, respectively, of person perception and of group dynamics. It is essential, however, to know where the topics fit in the justice system if we are to form sound judgments as to the relevance of what social psychologists have done, and so Chapter 7 begins with an overview of the system.